Usability Assessment of a Powered Wheelchair Controller:

How Impairments Affect Human Computer Interaction Based Tasks

---

A thesis submitted in partial fulfilment of the requirements for the Degree of

Masters in Engineering

in the

University of Canterbury

by

Rory M. Horne

University of Canterbury

2014

Supervisor - Dr Dirk Pons, College of Engineering, Department of Mechanical Engineering

Co-Supervisor – Prof. William Helton, College of Science, Department of Psychology

---

Acknowledgements

I would like to thank my parents for always supporting me, my supervisor Dr Dirk Pons and co-supervisor Prof. Deak Helton for their guidance and advice, Dynamic Controls whom I worked with and the University of Canterbury for putting me through the ups and downs of an engineering degree. Additionally I would like to thank Roxanne and Pat for their help proof reading.

Finally I would like to thank You, the person who is actually reading someone's thesis.

Abstract

Problem: Designing the user experience is a growing trend in product design; however this trend has not greatly benefited people with impairments and disabilities. There are no practical tools to broadly assist with this issue. There is a need for standardized measures to quantify impairment, a model to predict how designs may perform and a need for data regarding how people with impairments interact with consumer technology.

Purpose: To conduct a usability analysis with an industry partner on their powered wheelchair controller using participants with varying impairments. The industry partner was seeking better insight into the benefits of formal user testing.

Method: Forty consenting adults were given a score representing their level of impairment using six measures from the International Classification of Functioning (ICF). These measures were identified by the researcher to affect interaction with a device. Performance was measured by time taken to complete tasks, errors made, reported task difficulty and reported controller usability.

Results: Performance was reduced in participants with a higher ICF score and age. An ICF score less than or equal to 2 was 117 times more likely to not complete the tasks, greater than or equal to 3 was not able to complete the experiment. Age >50 years took an average 79 seconds longer than <35 years to complete a task and reported greater difficulty, more errors and a lower usability for the controller.

Implications: Low to moderate levels of impairment has a significantly negative effect on the usability of common devices. Difficulties were mostly cognitive with participants unable to create an accurate mental model of the system. Participants with lower performance tended to be overly optimistic about their abilities. Mistakes were the greatest source of error followed by lapses and almost no reported or observed slip errors.

Original Contribution: The ICF has never been used as a metric for usability testing. This study successfully applied the ICF alongside other measures to prove its validity. Based on the results and current literature the Task Process Model was created to provide a simple and practical way to describe the interaction of people completing a task of basic to moderate complexity.

Author Contact: rory.m.horne@gmail.com

Contents

Table of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 6

Abbreviations and Glossary

<35	Participants grouped together for the purpose of data analysis who were under 35 years. There were 15 participants in this group.
50+	Participants grouped together for the purpose of data analysis data analysis who were over 50 years. There were 12 participants in this group.
DNC	Participants grouped together for data analysis purposes who did not complete (DNC) all four tasks, that is they asked to stop the experiment or were judged to be unable to complete a task at some point in the experiment. There were 13 participants in this group.
Disability	A complex interaction between features of a person's body and features of the environment and society in which the person lives. This is an umbrella term, covering impairments, activity limitations, and participation restrictions (Rosenbaum & Stewart, 2004).
HCI	Human Computer Interaction
ICF	International Classification of Functioning, Disability and Health. A framework created by the World Health Organisations for measuring health and disability at both individual and population levels.
ICF specifics	The individual six measures of the ICF assessment.
Impairment	A problem in body function or structure (Rosenbaum & Stewart, 2004)

GUI	Graphical User Interface
Performance Time	The time taken in seconds for a participant to complete each attempt at a task.
PsycMotor Functions	Term used for the ICF measure for 'quality of psychomotor functions'
SUS	The averaged System Usability Score for a participant
SUS Specifics	The individual ten components of the System Usability Score. (these ten components are described below)
SUS Liking of System	I think that I would like to use this system frequently
SUS System Complexity	I found the system unnecessarily complex
SUS Ease of Use	I thought the system was easy to use
SUS Support for Use	I think that I would need the support of a technical person to be able to use this system
SUS Function Integration	I found the various functions in this system were well integrated
SUS Inconsistency	I thought there was too much inconsistency in this system
SUS Speed to Learn	I would imagine that most people would learn to use this system very quickly
SUS Cumbersome	I found the system very cumbersome to use
SUS Confidence in use	I felt very confident using the system.
SUS Amount to Learn	I needed to learn a lot of things before I could get going with this system.
Task 1 Chair	Required participants to activate the feature on the controller that would raise the chair. As with all tasks participants are asked to attempt the task four times consecutively.
Task 2 Lights	Required participants to turn the light feature to 'on'. As with all tasks participants are asked to attempt the task four times consecutively.

Task 3 Clock	Required participants to turn the clock feature to 'on'. As with all tasks participants are asked to attempt the task four times consecutively.
Task 4 Cog	Required participants to perform the first three tasks while being cognitively loaded with a word association test. Participants are asked to attempt the task four times consecutively
TLX End Demand	This value is an average on mental demand, physical demand, temporal demand, frustration and effort. The self reported demand of a participant from the NASA TLX (Task Load Index) assessment upon completion of the four attempts as a single task.
TLX Start Demand	As with TLX End Demand but upon completion of the first attempt of a single task.
TLX End Performance	The self reported performance of a participant from the NASA TLX assessment upon completion of the four attempts as a single task.
TLX Start Performance	As with TLX End Performance but upon completion of the first attempt of a single task.
TLX Specifics	The individual six measures of the NASA TLX assessment.
Total ICF	The summed points of ICF assessment for a participant for all ICF measures.
TPM	Task Process Model. The concept model created in this study to describe the process of complaining a task including external and internal factors that affect the outcome.
Voluntary Movements	Term used for the ICF measure for 'control of simple voluntary movements'

1.0 Introduction

There is a growing emphasis on better usability based design in industry and a deficiency in how well technology meets the usability needs of people with impairments.

1.0 Importance of User Centered Design

''The need for the future is not so much computer oriented people as for people oriented computers'' (Nickerson, 1969, p. 178)

The quote given above highlights a technology gap that exists. 'People oriented computers' are designed for the mean of the normally distributed population. Those remaining on the outskirts are being left behind to use adaptations of outdated consumer electronics, unable to use the centre focused technology which continues to improve at an exponential rate.

An example of this technology lag is that even Stephen Hawking recently received new hardware and software to aid his struggle with motor neuron disease that has been superseded by newer editions for at least one year ("Stephen Hawking gets a tech upgrade from Intel, but keeps his original voice | News | Geek.com," 2014, "ThinkPad X Series," 2014, "Windows 7 editions," 2014).

Computers are involved in almost every facet of the modern western lifestyle with User Centered Design focused more on Human Computer Interaction (Ritter, Baxter, & Churchill, 2014). Human Computer Interaction (HCI), as its name suggests focuses on how humans interact with computers. A large part of this focus is on the visual display of the computer, physical input by the user and information communication between the user and the computer.

User Centered Design, Human Factors, Ergonomics, HCI, Usability Design and User Experience Design are all based broadly on allowing technology to be used with greater effectiveness, efficiency and satisfaction. This is done by giving weight to the end users of the product as the driver for design. The constraints and goals of the product design are based on the constraints and goals of the user.

Due to the increasing penetration of computers into daily life and their involvement with a variety of activities there is greater necessity to provide satisfaction on top of effectiveness and efficiency. This is because for many people interaction with computers is a pastime activity, such as video gaming, online shopping and use of the Internet for media streaming. An example is the rise of casual games (low cost, wide spread and simple digital games) where an estimated 59% of 6 to 65 year olds in the United Kingdom play some form of video game (Juul, 2012).

To succeed in the current market place products need to not only provide functionality and efficiency but also an enjoyable experience for the user. For example a pleasurable experience is paramount to success in the current marketplace for mobile apps as they are easily available to users, obtainable for very little cost and many functionally similar products exist,.

1.2 Ability and Disability

Consumer electronics are being improved upon constantly and new methods of interacting with technology are continually conceptualised, developed and realised to market. Touch screens are common place, motion control has been integrated into video gaming, voice recognition is widely used on cell phones, 3D movies and television are a current trend, and a new wave of virtual reality and immersion technology is approaching quickly on the horizon.

These consumer goods predominantly cater for able bodied users. These users are assumed to meet particular levels of physical ability, sensory perception, cognitive processing and experience with technology. For people with impairments using many consumer electronics is difficult or impossible.

For those people requiring specialist devices to aid them these products lag behind the improvements and developments of more common place consumer electronics. The manufacture of specialist devices for aid does not benefit from the scales of economy to reduce the cost of research and development as is the case for many consumer manufacturers.

In addition current consumer technology development benefits from the wealth of community discussion, reviews and feedback given directly to the relevant designers and manufactures through the internet.

Products catering for people with impairments often have a less direct link with the end users. A typical business model involves the design and manufacturing companies working with larger sales firms through a tender for the rights to supply a government department, who in turn rely on medical professionals to prescribe the required product for the end user. The end user may have very little choice of products and lack the channels to report any frustrations they have. This lack of feedback stagnates usability design for aid deceives, especially compared to common consumer electronics.

Gaining insight into how people with impairments actually use the aid devices is invaluable information for a number of reasons. Firstly it is rare to get such feedback, secondly due to the specific abilities of the users it is a poor comparison for an able bodied person to represent them and thirdly there is likely to be a large difference between how the aid device is intended to be used and how it is actually used.

There is a need for better understanding of how specific impairments affect usability in order to create greater access to current technology for those requiring specialist aid devices. Having a better understanding of how impairments affect usability will allow designers to focus their attention on which aspects of a device are most critical for accessibility.

1.3 Specific Industry Situation Examined

This thesis project involved working with an industry partner, a company which designs and manufactures controllers for power wheelchairs. This partner has a new generation of integrated controllers that are able to link with consumer electronics such as smart phones, tablets and personal computers. This has emphasised for them the importance of usability in product design.

The industry partner wanted to gain a better understanding of what aspects of the controller were the least usable and for which type of user. This included which types of errors were made and the user's emotive response to using the controller. It was also important to represent a common use environment by including distractions (cognitive loading).

1.3.1 Wheelchair Controller Analysis

The analysis was based around the powered wheel chair controller seen in Figure 1.1. Although the controller has now been replaced, the physical controller and base menu structure remains the same for the current iteration. The usability analysis of the controller will not be on its driving and handling controls, only on its secondary features such as adjusting the seat. It is noted that the controller was not attached to a wheelchair but mounted on a stand.

The partner had no formal in house usability or user experience analysis methods to expand on; consequently the direction of the study was unconstrained and flexible. A primarily quantitative usability analysis was decided upon as numeric data is more reliable and less open to interpretation.

This study was worth conducting as the results may be used to influence the design of future products rather than the controller being tested for the industry partner. For this reason the analysis is focused on understanding the users as opposed to the specifics of the controller.

The participants were an integral aspect of the project, requiring a diverse range of abilities from able to disabled to allow for comparison. A total of forty participants were found that represent old and young as well as a range of ability. Participants were sourced through various care centers, disabled support networks as well as through advertising.

Figure 1.1: The wheelchair controller assessed in this study

1.4 Purposes

There were four major objectives of this study with the intention to both assist the industry partner directly and to give novel contributions to the associated bodies of knowledge. This study will contribute to the fields of usability engineering and user experience analysis, these areas are the intersection of engineering product design and psychology. The study purposes reflect this unique field as they center on converting qualitative information into practical data.

The study purposes are:

1. *Identify the effects of impairments on a user's measured performance, and self-reported performance and usability experience.*

The main focus of this study was to better understand what aspects of the wheelchair controller are the least usable for people with particular impairments. Performance here refers to the measured learning curve of the participants as well as self-reported performance. It is likely that other variables will have a sizeable impact on usability in particular ages. Understanding how impairments and other factors affect the usability of the wheelchair control will allow for extrapolation to how usability on other devices may be affected.

2. *Determine how cognitive loading affects a user's measured and reported performance.*

Here cognitive loading refers to something demanding attention that is not associated with operating the controller. Understanding the effects of cognitive loading is important not only to represent use in 'the real' world but because different impairments combined with external distractions may cause unexpected results in performance.

3. *Identify what types of errors are made when using the controller and the possible cause for these.*

Classifying error types will help to determine what specific difficulties users are having with the controller. Errors will be classed in a broad context as opposed to identifying errors specific to the controller such as the likelihood of pressing one button over another.

4. *Create a model that can be used to explain the influence of the measured variables.*

Representing the study findings in a visual flow chart style model will allow for better and easier understanding of the results and allow broader and easier practical application in a design setting. The model will represent the process of a user deciding upon and performing an action and how impairments and other factors may have an effect. The model will be intended for general use by a wide range of practitioners.

1.5 Preparation for Study

To gain relevant knowledge and experience the researcher enrolled and observed several university courses, as well as special topic and independent studies relating to usability engineering and HCI before beginning the usability assessment project.

The courses included Training and Learning in the Work Place, Environmental Psychology, and Human Factors/Ergonomics. All of which were run through the University of Canterbury Psychology Department, specifically in the postgraduate program of Applied Psychology.

The special topic and independent studies involved self-directed research resulting in a final report on the subject. The studies undertaken were Group Decision Making, Human Error, and Performance in Engineering Teams.

The Engineering Teams study resulted in a journal paper being published tilted 'Industry Based Team Projects: Personality Traits That Influence Success in Engineering Education'[1]. The full journal paper can be seen in Appendix 1.

This paper involved many of the same statistical techniques used in this study as well as providing experience with the statistical analysis software SPSS. In addition a conceptual model was created describing the various factors affecting engineering team performance which provided lead the way for the design of the model created in this model.

[1] Industry Based Team Projects: Personality traits that influence success in engineering education. Horne and Pons (2012) Journal of Adult Learning Aotearoa New Zealand (JALANZ), Vol. 40, No.1, 50-71

2.0 Literature Review

The literature review presented here is intended to provide context for the study by summarising the history and current state of User Centered Design, Human Computer Interaction (HCI) and Error analysis. More specific to this study the history and current methods of quantifying impairments and disability are reviewed.

A variety of information processing models are also reviewed, as a purpose of this study is to create a novel model for simple explanation and application of these results.

Finally gaps in the body of knowledge are identified.

A review of methods used for usability analysis can be found in Appendix 2 which identifies tools that were both appropriate and impractical for use in this study.

2.1 User Centered Design

The concept of making a task, device or system to be best used by a person for greater productivity is not a new. Ergonomics emerged as a scientific discipline in the 1940s, predominantly from the military. This created a growing realisation that as technical equipment became increasingly complex not all of the expected benefits would be delivered if people were unable to understand and use the equipment to its full potential ("A brief history of ergonomics and user centered design," 2014).

Ergonomics (or human factors) is defined as the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance ("Definition and Domains of ergonomics | IEA Website," 2014).

The common use of the term 'user interface design' has evolved from invoking images of tangible levers and dials placed in arms reach by engineers into the design of web pages and mobile apps made on illustrator software by graphic designers.

User Centred Design (UCD) came to the fore during the 1980's with a broader focus than previous approaches. It began to look at the user's needs, carrying out an activity/task analysis as well as a general requirements analysis, carrying out early testing and evaluation, and designing iteratively (Ritter et al., 2014).

Human Computer Interaction (HCI) is considered to be in its third wave. Moving from the first and second wave which focused on purposeful tasks within work place settings. The third wave focuses on the use of computers in private and public settings as means to have cultural and emotive experiences (Bødker, 2006). This third wave of HCI falls under the greater umbrella of UCD and has influenced the rise of the User Experience (UX) field. UX embraces the reality that computers now come in many forms and influence people on many levels, with effects beginning, ending and tying in with more than just the screen (Hassenzahl, 2008; Hassenzahl & Tractinsky, 2006).

The UX field has been growing steadily (Hassenzahl & Tractinsky, 2006) and currently the demand for UX jobs is increasing (Beecher, 2014; "Technology Vision 2014: Key IT Trends - Accenture," 2014).

UX can be defined as "a consequence of a user's internal state (predispositions, expectations, needs, motivation, mood), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use)" (Hassenzahl, 2008, p. 6).

This definition of UX aligns with the third wave ideas of HCI, both emphasize that a person's interaction with a system is not restricted to simply how that interaction occurs but that many factors will affect how that interaction happens. This includes responses on many levels from the user including an emotive one. Human Factors and Ergonomics and to a greater extent UX have grown to include many aspects of psychology as well as engineering.

2.2 Key Aspects of HCI

To create a simple and practical means to improve interaction between a person and a piece of technology, both the operator's ability and the usability of the technology need to be quantified.

2.2.2 Current Three Key Aspects of HCI

HCI is arguably the most important aspect of usability in the current age. Computers are involved with most aspects of life in the developed world and people are interacting with computers in increasingly different ways.

Based on the following literature three key aspects of HCI have been identified by the researcher. Shown in hierarchical order these are:

1. Functionality of the system – Is the system capable of being used to do the tasks for which it was intended for?
2. Users ability – Does the user have the physical and cognitive ability to operate the system?
3. Communication of the system model to the user – Does the system provide the users, who have the ability to operate it, with instructions, cues and affordances to effectively use the system?

Functionality of the system refers to whether the system is usable at its most basic level. That is, does the software and hardware work as they were designed to, independent of how a user would rate the products usability. This aspect would be controlled mostly by the engineers during product manufacture.

A product is likely to be designed with particular users in mind independent of whether user centered design or user experience design principles are being used. Once a product is considered functional enough for use it needs to be considered whether the intended user and others have the ability to operate the product.

After it is determined that users have the ability to operate a product, the effectiveness of communication between the two can be considered. This relates to the affective aspects of HCI, if there is poor communication the user will have a negative emotional response towards the product.

2.2.1 HCI as Chronologically Viewed by the Literature

"Eventually a great variety of mental processes should be interpretable directly by electrical circuits." Said by Page (1962), speaking on (HCI) as he believed it would be in 2012. When there were still less than 20,000 computers in existence thought was being put into how HCI could one day be made seamless (Freed & Ishida, 1995).

The modern world has the technology to measure brain function in MRI machines, control robotic prosthetics from the electrical signals in muscles and read brainwaves to allow basic control over computers. However the gap between our mental processes and the control of technology is still predominantly bridged by an interface requiring inputs and outputs based on our physical senses and abilities.

The two major areas identified for effective human computer interaction coupling were man-to-machine and machine-to-man (Page 1962) . That is:

- Machines must respond to forms of communication from man
- Man must be able to draw information from the machine

It was idealized that the machine interface would begin to dissolve completely describing a man-to-man interface evolving from the man-to-machine interface. Currently the man-to-machine interface or user interface as it is now referred to still pose many constraints that prevent the seamless cognitive links between people as envisaged by Page.

As computers became more common one of the first considerations of user interface design was for time-sharing systems (sharing a computing resource among many users); the need for time-sharing required multiple users with different set-up requirements to use a single computer (Cheriton, 1976). The different needs of various users lead the computer interface to be described as creating the environment in which the user interacts with the system (Cheriton, 1976). From this environment the user creates a mental model of the system, which they used to predict the systems' behavior, this environment consists of several concepts put forward by Cheriton (1976):

- A universe of objects and actions - the user specifies actions to be performed on objects.
- A command language - the language in which the user specifies these actions and objects.
- A system response language - the language in which the system responds to user commands
- The dialogue structure - the structure underlying the man to computer dialogue determining the behaviour of the system.

In 1991 the year that the Macintosh PowerBook series was released the fundamental aspects of HCI had changed within the literature to recognize that people differed in their ability to physically interact with computers and cognitively grasp that interaction (Downton, 1991; Freed & Ishida, 1995).

The model of a computer was described as having a central processor with inputs and outputs for communication with the outside world. For a person to interact with a computer in relation to these three components specified above, the user must possess the following (Downton, 1991):

- Senses – To recognize the outputs of the computer, namely vision to see graphics, hearing to listen to audio indicators and touch to feel for kinetic feedback.
- Motor control – To give inputs to the computer, for example through a keyboard and mouse, movement or the eye for tracking (Countering this brain wave monitoring to give commands is possible, although it is not yet commonplace).
- Cognitive ability – To understand the required inputs and outputs and create a mental model of the system.

Three different levels of user activity identified based on previous research were used to describe HCI; these are physical, cognitive and affective (Karray, Alemzadeh, Saleh, & Arab, 2008).

- Physical – Mechanics of interaction between human and computer
- Cognitive – Ways in which the user can understand the system
- Affective – The user relation to the computer on an emotional level

Affective computing was first coined by Picard (1995). This idea of the relationship between the user and the computer on an emotional level, was popularized by Picard (2000) in her book Affective Computing. The term Affective Computing arose in reference to Page (1962) describing the affective aspects of human communication, similar to the more subtle and emotionally based communication between people such as gestures, eye movement, tone of voice and expression.

Throughout the development of HCI it has been clear that the efficiency of any interaction between humans and computers is based on various forms of communication. Computers as designed devices are the party in this communication exchange that are most able to be adapted to people, rather than people being forced to work around the constraints of computers. To reduce errors in HCI an understanding is needed of what errors types of errors people make.

2.3 Classifying and Identifying Error

This section of the literature review focuses on how human error is classed and identified. It begins with a review of work done by Jen Rasmussen followed by James Reason, two of the most influential contributors to the study of human error. James Reason defined the error classes of Slips, Lapses and Mistakes. Definitions are given below for these classes appropriate to this study.

A brief overview of the modern approach to human error is given, in that the term is becoming obsolete and human error is seen only as a link in system failure rather than a sole cause.

Finally methods are explored to identify errors in other usability studies.

2.3.1 Framework of Skills, Rules and Knowledge

Jens Rasmussen explored the idea of human error at the machine interface and its effect on developing digital technologies (G. I. Johnson, Clegg, & Ravden, 1989; J. Rasmussen, 1983) followed by application to large scale industry accidents (Jens Rasmussen, 1982, 1997; Vicente & Rasmussen, 1990).

Jens Rasmussen (1982) created a skill-rule-knowledge based framework originating from a verbal protocol study of technicians engaged in electronic troubleshooting. The framework is a tripartite distinction of performance levels. The performance level of human operators can be classified as skill-based, rule-based or knowledge-based.

Behavior at the skill-based level represents a sensory input to motor output, which occurs without conscious control, where the motor output is a smooth and automated pattern of behavior. Generally skill-based performance occurs in much quicker time frames than rule-based or knowledge-based performance; smooth instinctive actions do not require long periods of planning immediately before they are performed (Rasmussen, 1983).

When following preset rules to solve familiar problems, behaviour is considered to be rule-based. Examples of rule-based behaviour are following a recipe from a cookbook or performing standardised quality control methods (Rasmussen, 1983).

Knowledge-based behaviour requires the greatest level of conscious thought in Rasmussen's framework. This level of performance is used in novel situations where no learned responses or designated rules from previous encounters can be applied (Rasmussen, 1983). Performance must be moved to higher conceptual level where action must be planned and thought through using analytical processes and stored knowledge (Reason, 1990a, 1990b). It can also be assumed that the majority of knowledge-based performance has fewer time constraints than other levels of performance.

2.3.2 Classes of Slips, Lapses and Mistakes

Reason developed the Generic Error-Modelling System (GEMS) to locate the origins of practical human error (Reason, 1990a). This system is based closely on the performance level framework created by Rasmussen (1983). In addition a variety of works from various cognitive processes studies have contributed to the GEMS including attention and mental resource theories; models of memory and rationality; and cognitive control models (Reason, 1990a).

The three basic error types outlined by Reason are skill-based slips and lapses, rule-based mistakes and knowledge based mistakes. Rasmussen's levels of performance are distinguished by the level of cognitive thought required. Reason carries this to his classes of error by asking 'whether or not an individual is engaged in problem solving at the time when an error occurred' to link the error classes to the performance levels (Reason, 1990a, p. 56).

The associated performance levels and error types are seen in Table 2.1. In addition to building on Rasmussen's work Reason's work is based closely on the differentiation of slips and mistakes first suggested by Norman (1988).

Table 2.1: Performance levels and error types

Rasmussen's Performance Levels linked with Reason's Error Types	
Performance Level	**Error Type**
Skill-based level	Slips and lapses (mode errors)
Rule-based level	Rule-based mistakes
Knowledge-based level	Knowledge-based mistakes

Error is used by Reason as "a generic term that encompasses all occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some chance agency" (Reason, 1990a, p. 9).

A graphical interpretation of slips, lapses and mistakes occurring in the execution of an action can be seen in Figure 2.1. This figure shows when the results of the three error classes become apparent, not when the errors are committed.

The performance levels by Rasmussen and error types identified by Reason are still practically applied in the modern research for improving systems and product design, however their application appears to be uncommon beyond research and specialist industry (Ferner & Aronson, 2006; Glavin, 2010; Li, 2011; Lopez, Love, Edwards, & Davis, 2010; Mattox, 2012; N. A. Stanton & Salmon, 2009). This may due to their abstract nature and the difficulty in discriminating the error type from a generic error. Incorporating these taxonomies into a forward planning tool may allow them to be better understood better by the initial designer rather than just the product tester and reliability specialist.

The definition of slips, lapse and mistakes are as follows based on (Boff, 2006; Helander 2006, pp. 337-338; de Lange & van Knippenberg, 2009; Reason, 1995, 2000):

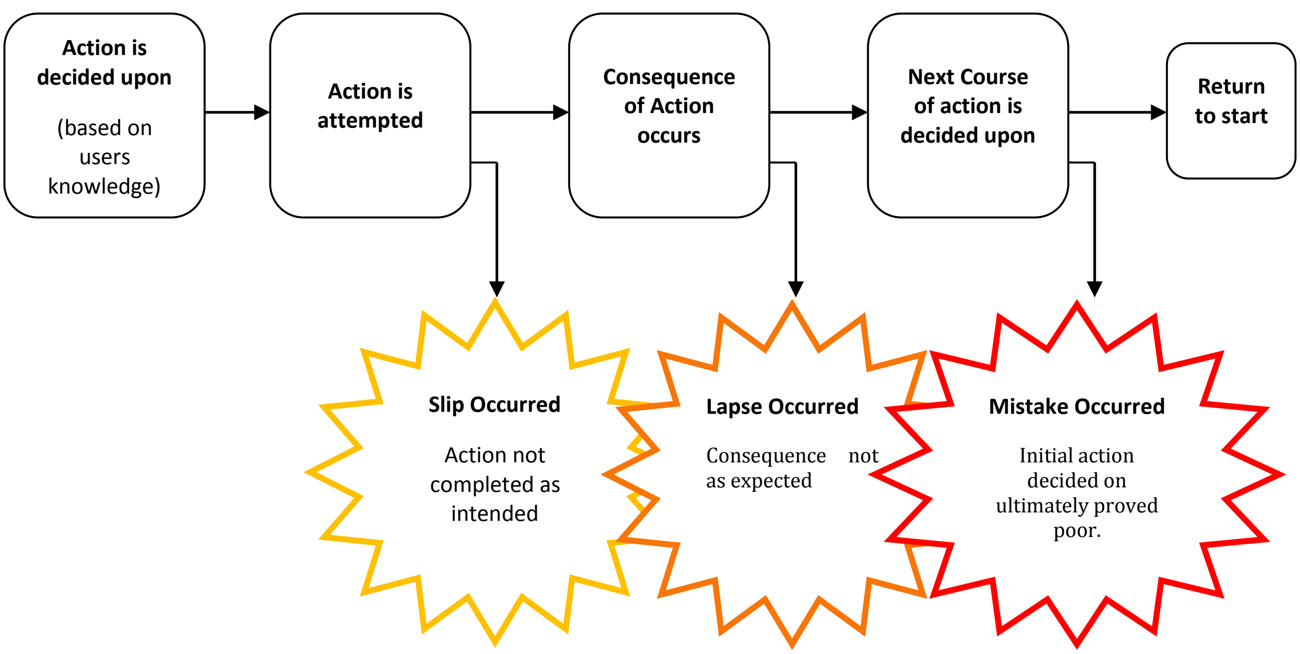


Figure 2.1: Occurrence of Reasons errors in a performed action sequence (Image by R Horne)

Slip

A Slip is an external action (as opposed to an internal action such as a decision being made) that is potentially observable, such as a misspoken word or spelling mistake. Slips occur during skill-based performance and result from some failure during the execution of an action sequence.

Slips occur regardless of whether or not the intended action was adequate to achieve its objective, if no slip were made a Mistake may still have occurred.

Lapse

Like a Slip, Lapses occur during skill-based performance but unlike Slips they occur when the mental recall of how an action is performed is done poorly, mostly due to distraction or impairment (temporary or permanent).

A lapse is a more covert form of error compared to a slip and is not likely to manifest as an obvious erroneous action. Lapses, more so than slips, may go unnoticed for a significant period of time after the action is complete, depending on whether the stakeholder involved views the action as successful.

Lapses can also be referred to as mode errors, the operator may forget (have a lapse in memory) of the mode in which they are operating in. For example a driver who is tired, distracted, nervous, drunk or inexperienced may put the car in reverse 'mode' when performing a three point turn rather than drive 'mode' from neutral, recalling incorrectly which direction to move the gear stick (not which direction they want the car to go). This demonstrates how a lapse is only noticed after it occurs due to the desired result not occurring later in the planned process. A slip, such as missing the gear stick, is clearly observable, as the planned process cannot continue without the stick being grasped.

Mistake

Mistakes are defined by Reason as "deficiencies or failures in the judgmental and/or inferential processes involved in the selection of an objective or in the specification of the means to achieve it, irrespective of whether or not the actions directed by this decision-scheme run according to plan." (Reason, 1990, pg. 9).

Mistakes can take the form of either rule based or knowledge based. Knowledge-based mistakes occur when a user poorly understands and interprets the situation (due to a lack of knowledge). The decisions made and resulting motor actions from these decisions are unlikely to give the desired outcome. Knowledge-based mistakes are usually obvious to the user as they are aware that they are confused, they may be observable due to the hesitancy and visible confusion of the user.

Rule-based mistakes occur at a less cognitively intense level for the user. Similar to a lapse they occur when the user has issues recalling previous information relating to the situation. A Rule-based Mistake occurs when the user has recalled multiple courses of action that may apply to the current situation and a decision must be made about which course of action to take.

2.3.3 Modern Concept of Error

The term Human Error is becoming out dated and even inappropriate. "It [the label of human error] retards rather than advances our understanding of how complex systems fail and the role of human practitioners in both successful and unsuccessful operations." (Dekker, Cook, Johannesen, Sarter, & Woods, 2010, p. 4).

Current understanding is that an erroneous human action is one link in a chain of events involving both mechanical and managerial errors, all culminating in an accident (C. Johnson, 1999). The environment created by the organisational structure, as well as by other factors such as mechanical design and operator training all contribute towards a final result, as opposed to the fault lying solely with the operator's action (Dekker et al., 2010; Helander, 2005).

As is the case with this study, it may be the participants who ultimately make the error but that erroneous action is only the point of a wedge made of many other factors (Reason, 2000). The minimal training to use the controller, previous experience of the user operating similar devices, as well as multiple aspects of the controller's design including physical layout, icon clarity and general intuitiveness all add to the likelihood that a participant will make some form of error.

2.3.4 Human Error Identification Methods

There have been many methods developed for Human Error Identification (HEI) since the 1980's that were influenced by Rasmussen (1982) and later by Reason (1990). Noted HEI methods included Systematic Human Error Reduction and Prediction Approach or SHERPA (Embrey, 1986), Generic Error Modelling System or GEMS (Reason, 1990a), Cognitive Reliability and Error Analysis or CREAM (Hollnagel, 1998; Marseguerra, Zio, & Librizzi, 2006), Human Error Identification in Systems Tool or HEIST (B. Kirwan, 1994) and Task Analysis For Error Identification or TAFEI (N. A. Stanton & Baber, 2002) to name a few .

Kirwan conducted many reviews and comparisons of HEI techniques to determine what defines an effective method and which of the tested methods are effective (Kirwan, 1992a, 1992b, 1998a, 1998b; Shorrock & Kirwan, 2002). The need for contextual validity was particularly important to this study as it deals with a range or user abilities although none of the methods reviewed were suited for direct application to this study.

Many of the HEI techniques are based on observation of the task, post task user interviews and indirect association of errors and other measures. Despite considerable work in the areas of HEI it has been suggested that methods such as the ones stated above have had little practical impact in industry (C. Johnson, 1999; Lucas, 2001).

2.4 Cognitive Loading

Of the three errors types discussed above (slips, lapses and mistakes) external distraction affecting the practitioner may be the cause of, or reason for these errors. The chances of lapses and mistakes occurring are particularly sensitive to the effects of distraction and reduced cognitive processing (Chandler & Sweller, 1991; Sweller & Chandler, 1991). Distraction and reduced cognitive processing are described by the theory of cognitive loading.

Cognitive loading is the theory that during complex learning activities the amount of information and interactions that must be processed simultaneously can either underload or overload the finite amount of working memory one possesses (Paas, Renkl, & Sweller, 2004; Sweller & Chandler, 1991).

Cognitive loading can be divided into three types, visually represented in Figure 2.2 (Chandler & Sweller, 1991; Sweller, van Merrienboer, & Paas, 1998):

Intrinsic

The inherent level of cognitive load associated with a specific instructional topic. In the case of this study the intrinsic cognitive load would be the inherent difficulty to perform and understand each task in the experiment and to comprehend the menu structure and basic physical operation of the controller.

Extraneous

The cognitive load generated by the manner in which information is presented to learners. This load is under the control of the designers. In this study minimal information is provided to participants. This restricts the extraneous load, related to the understanding of the controllers operation, mostly to the design of the controller. In this study this is the cognitive loading induced by the researcher through explanation of how to use the controller.

Germane

The cognitive load is that load devoted to the processing, construction and automation of schemas, that is, the conversion of sensory input into organised knowledge, information, categories and rules. In this study the germane cognitive load is effectively an independent variable i.e. the cognitive abilities of the participants, which will likely vary with impairment and other factors.

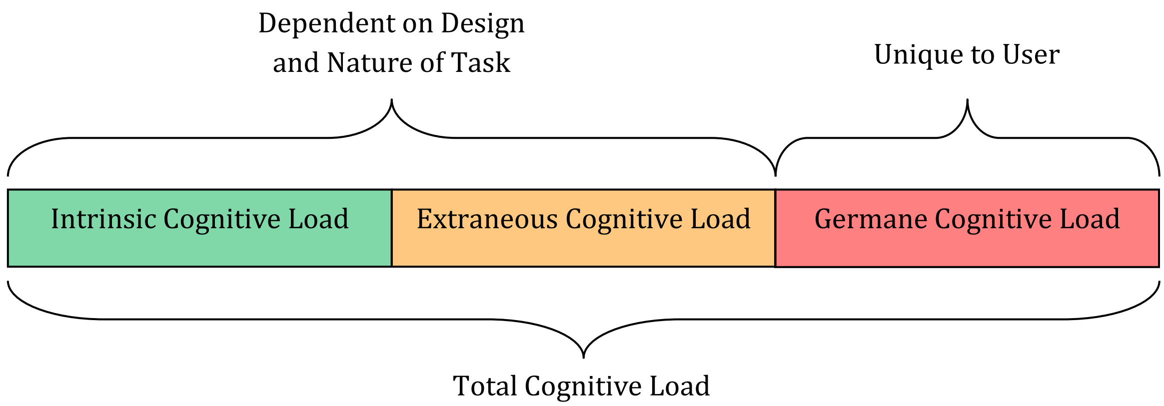


Figure 2.2: Scale of cognitive loads (Image by R Horne)

Cognitive loading is often used to represent distractions that may occur in real world situations (Engström, Johansson, & Östlund, 2005; Erk, Abler, & Walter, 2006; Goodell, Cao, & Schwaitzberg, 2006; Lam, 2002; Liang & Lee, 2010; Ward & Mann, 2000).

2.5 Quantifying Impairment and Disability

The next issue to consider after understanding the nature of HCI and error is that the physical and cognitive abilities of people differ with respect to the requirements for interaction with a computer based device. Specifically it needs to be understood how the abilities of users can be quantified with regard to the interfaces and specifically the wheelchair controller in this project.

The first subsection summarizes recent models of disability. These models identify views of disability treatment, how disability is accepted by society and where the 'problem' of disability arises.

The second subsection describers the International Classification of Functioning based on the latest and most accepted model of disability.

2.5.1 Models of disability

Although a large number of models have been identified the medical, social and the current social adapted model are explored as they directly impact directly the ability quantification tools used in this study.

Other models have influenced the development of these models and provide insight into society's changing views on the treatment of the disabled; displaying a range of attitudes from caring to denial, to damning. However details of these models are superfluous to the final outcome of this study.

The Medical Model of Disability

The medical model of disability was most prominent in 20th century western medicine (Boorse, 1975, 1977). It became common with the publishing of the International Classification of Impairments, Disabilities and Handicaps (ICIDH) by WHO in 1980 which contained information on diagnosis and health condition but not functional status or overall well being of the individual (Bickenbach, Chatterji, Badley, & Üstün, 1999; De Kleijn-De Vrankrijker, 2003; Simeonsson, Lollar, Hollowell, & Adams, 2000).

The Medical Model holds that disability results from an individual person's physical or mental limitations, placing the source of the problem within the impaired person and concludes that solutions are found by focusing on the individual (Brisenden, 1986; World Health Organization, 1980). The general view in the medical model is that disability is a curable, treatable or containable condition that is largely unconnected to their social or environmental context (Brisenden, 1986; G. L. Engel, 1977).

The model has been criticised for the abnormalisation of disabled people and imposing a paternalistic approach to problem solving which, although well intentioned, concentrates on "care" and ultimately provides justification for institutionalisation and segregation (Llewellyn & Hogan, 2000).

Another criticism is that the medical model fostered existing prejudices in the minds of employers. Because the conditional is "medical", a disabled person is seen to be prone to ill health and sick leave, is likely to deteriorate, and will be less productive than work colleagues (Llewellyn & Hogan, 2000).

The Social Model of Disability

The social model of disability arose as a reaction to the more dominant medical model by organisations and communities other than WHO (Paley, 2002; Segregation, 1976). The comparison between the two models may be described as an attitude change from "cure to care" (Nikora, Karapu, Hickey, & Te Awekotuku, 2004, p. 6). The phrase 'social model of disability' was coined by Mike Oliver in 1983, an academic with a disability (Oliver & Sapey, 2006).

The Social Model views disability as a consequence of environmental, social and attitudinal barriers that prevents people with impairments from maximum participation in society (Davis, 2013). Disability is seen to stem from the failure of society to adjust to meet the needs and aspirations of a disabled minority and that society must adapt for the benefit of those with disabilities (Oliver & Sapey, 2006; Shakespeare & Watson, 2001). This model implies that the removal of attitudinal, physical and institutional barriers will improve the lives of disabled people, giving them the same opportunities as others on an equitable basis (Nussbaum, 2007; Shakespeare & Watson, 1997).

The strength of this model lies in placing the onus upon society and not upon the individual while still focusing on the needs of the individual (Shakespeare & Watson, 1997). The social model faces two main challenges (Goodley, 2001; Terzi, 2004), firstly, as the population gets older the number of people with impairments will rise, making it harder for society to adjust. Secondly, its concepts can be difficult to understand, particularly by dedicated professionals in the fields of charities and rehabilitation.

The Bio-psychosocial Model

The current International Classification of Functioning, Disability and Health (ICF) model is based on a combination of the Medical and Social Models of Disability, termed the Bio-psychosocial model by WHO (G. I. Johnson et al., 1989).

This model builds upon the Social Model, but incorporates elements of the medical model in that disability is both a problem at the level of a person's body, and also a complex and primarily social phenomenon. It accepts that impairments are significant physically, but stipulates that far more problems are created for disabled people by social and environmental causes (G. I. Johnson et al., 1989; "WHO | International Classification of Functioning, Disability and Health (ICF)," 2013).

The Bio-psychosocial Model, and hence the ICF, identify several contributing factors to a person's disability stemming from both health conditions and contextual factors (G. I. Johnson et al., 1989; "WHO | International Classification of Functioning, Disability and Health (ICF)," 2013), these factors are:

- Body Functions - physiological functions of body systems (including psychological functions).

- Body Structures - anatomical parts of the body such as organs, limbs and their components.
- Impairments - problems in body function or structure such as a significant deviation or loss.
- Activity - execution of a task or action by an individual.
- Participation - involvement in a life situation.
- Activity Limitations - difficulties an individual may have in executing activities.
- Participation Restrictions - problems an individual may experience in involvement with life situations.
- Environmental Factors - the physical, social and attitudinal environment in which people live and conduct their lives.

2.5.2 The International Classification of Functioning

The World Health Organisation International Classification of Functioning (ICF) Disability and Health known as the ICF was chosen to quantify the impairments of the participants. The ICF at the time of writing was in its tenth iteration and is designed for simple and practical use ("WHO | International Classification of Functioning, Disability and Health (ICF)," 2013). The ICF provides simple communication of the effects of impairments on task performance for both researchers and participants (Perenboom & Chorus, 2003; Reed et al., 2005), has proven application in measuring the effects of impairments on simple common tasks (Rentsch et al., 2003; Rosenbaum & Stewart, 2004) and is used in a wide variety of industry applications (Bruyère, Van Looy, & Peterson, 2005; Mortenson, Miller, & Auger, 2008; üStüN, Chatterji, Bickenbach, Kostanjsek, & Schneider, 2003).

From the literature few examples of the ICF being employed in usability design scenario were found. It appeared that the ICF was used to classify those with specific and severe impairments, namely people with cerebral palsy (Almanji, Claire Davies, & Susan Stott, 2014) and poor eyesight (Vigo & Harper, 2013). Other papers used the ICF as a definition for disability rather than applying the ICF (Fridin & Belokopytov, 2014; Ojasalo, Seppälä, Suomalainen, & Moonen, 2010).

The ICF has many measures that fall into three categories:

- Body functions and structure
- Activities (related to tasks and actions by an individual) and participation (involvement in a life situation)
- Environmental factors

Within these three sections there are hundreds of measures, six were identified which best described a person's interaction with the wheelchair controller. These six measures were chosen by the researcher based on the key aspects of HCI (discussed in section 2.1.0) and in particular the components of HCI identified by Downton (1991), which are Senses Motor Control and Cognitive Ability. The six measurers chosen from the ICF linked with the three HCI component are as follow:

- Senses – Visual Perception, Tactile Perception and Audio Perception
- Motor control - Control of simple voluntary movements and Quality of psychomotor functions
- Cognitive Ability – Ability to acquire new basic and complex skills

2.6 Sociotechnical Interaction Models

There are many models intended to explain, predict and otherwise represent human behaviour and how people interact with the environment. These range from complex mathematical algorithms utilising fuzzy logic (Sugeno & Yasukawa, 1993), foundational studies proving simple concepts like behaviour is learned from the environment (Albert Bandura, 1977; A. Bandura, Ross, & Ross, 1961), to popularised concepts such as the Hierarchy of Needs (Maslow, 1970).

The following is a review of models that have been found to best emulate a number of these qualities. Note that the models are described as being applied to an operator, which is the person who is performing a task.

A simplistic model representing the process of a person interacting with the world is shown as seven stages of action, seen in Figure 2.3 adapted from Norman (1988). This model accurately shows how a person approaches and performs a task, it's failing is that it is a linear model, ignoring any factors affecting the process and glossing over how previous experiences and the current state of the person modify perception, planning and action.

The following models and the model created in this study are essentially a more complex multi-path version of this linear process.

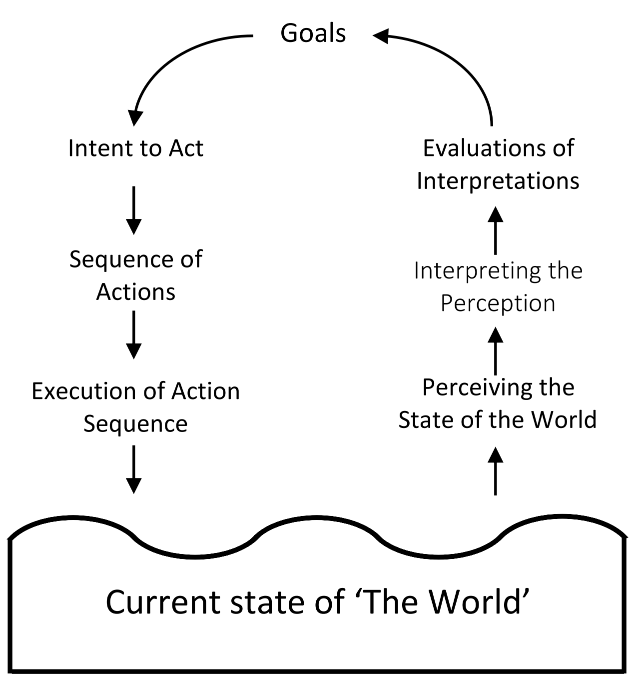


Figure 2.3: An operator's interaction with the world adapted from Norman (1988)

2.6.1 Information Processing Model

The model by Wickens & Hollands (2000) seen in Figure 2.4 describes a series of human information processing stages. It is intended to give a framework to analyse the different psychological processes an operator would progress through when interacting with a system to complete a task. This model is used as a basis to describe the relationship between the key concepts of observation, memory, response, attention as a resource and feedback.

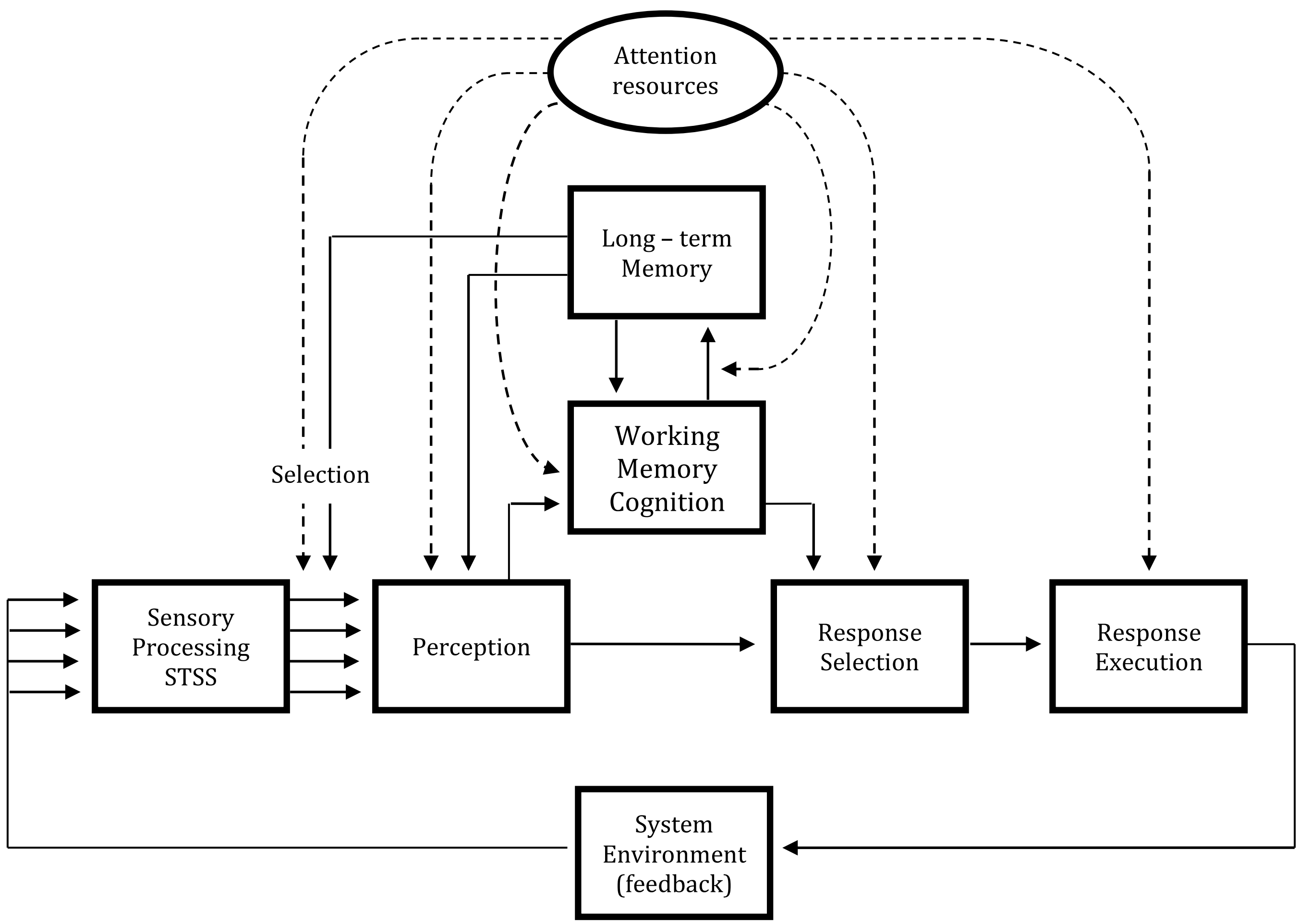


Figure 2.4: Human information processing adapted from Wickens & Holland (2000)

This model's major unique features are inclusion of an attention resource (D. A. Norman & Bobrow, 1975), long term memory, and working memory (A. Baddeley, 1992), and an environmental feedback loop which constantly monitors the operator's own motion and action as well as the motions and actions of other objects in the environment (Gibson, 1986; Warren & Wertheim, 1990). Like many of the models following there is a general flow from a sensory input to cognitive processing of this information and finally a physical output to perform an action.

In Wickens & Holland's (2000) model cognitive processing encompasses perception of the sensory information creating a loop through working and long-term memory. This loop feeds back to influence perception of the sensory information, leading onto response selection and finally leaving cognitive processing to create response execution. The information signals moving from the sensory processing to perception and onwards is termed bottom-up perceptual processing, while the feed loop providing information from long-term memory back to perception is termed top-down processing (Rumelhart, 1977). Bottom-up processing occurs when the sensory signals are able to provide significant information for the perception stage to assess the situation and move onto the response selection. Top-down processing occurs when sensory information is poor and assessment of the situation by perception requires additional information of similar previous experiences from memory to be able to move to a response selection (A. K. Engel, Fries, & Singer, 2001). Regardless of the situation and task there will always be a recall of previous experiences but the importance of this recall is determined by the strength of the sensory information at clarifying the ultimate response selection and execution.

The system environment feedback loop in Wickens & Holland's Model is used to identify two implications. First that there is a flow of information from the end of one task to the beginning of the next; that is a flow of information from the environment affected by the operator's actions, from which the operator receives sensory information and that their actions can influence the environment linking tasks. Second that there is also a continuous flow of information within the performance of a task, confirming to the operator what they are doing and what affect their actions are having. From this environmental feedback loop it can be seen that actions can create perception and that perception can create action (Neisser, 1976; Powers, 1973).

Throughout this process the attention resource, which is described as being finite, is drawn by the various stages of the cognitive process; therefore each stage becomes vulnerable to disruption and to error potentially as its supply of attention is diverted and reduced (D. A. Norman & Bobrow, 1975; Wickens & Hollands, 2000). Operators must use a mental strategy to effectively divide effectively and allocate their attention resource between the various stages of a task, different tasks and other mental operations occurring simultaneously (Kahneman, 1973;

Meyer & Kornblum, 1993). Errors can occur when the attention resource becomes scarce and therefore performance drops for one or more of the tasks or operations requiring it.

The Wickens & Holland model describes several stages of an operator's cognitive process when performing a task, capturing the dynamic flow between stages and how the process can become strained by the limitations of the attention resource. The model does not visually show the drive or motivation for the task, this is included in the sensory input and system environment feedback.

The model as a whole is predominantly cognitive, that is the majority of the stages are based in the operator's mind. Only the response execution and aspects of the system environmental feedback are physically observable stages, while other aspects of the model could be measured indirectly though biometrics. This makes the model impractical for experimental validation and potentially reduces its practical use as a measurement tool.

2.6.2 Internal Models

An internal model is a postulated neural process that simulates the response of the motor system in order to estimate the outcome of a motor command (Mitsuo Kawato, 1999; Wolpert, Ghahramani, & Jordan, 1995). The concept of an internal model became an important theoretical concept in motor control when it was determined that in order for voluntary movements to be controllable three computational problems needed to be solved (M. Kawato, Furukawa, & Suzuki, 1987):

- The determination of a desired trajectory in the visual coordinates
- The transformation of its coordinates to the body coordinates
- The generation of motor command

The three aspects of voluntary movement exist in three different levels, they are respectively a sensory recognition coupled with a desire as an information input, a cognitive process manipulating the input, and a physical motor output based on the manipulated input. Two classifications of internal models emulate the qualities sought for the model presented in this paper. These are the forward model and the inverse model.

Forward models, inverse models and their combination in systems are well explored topics in neuroscience and cognition and are well supported by recent behavioural studies in the field of sensory motor control (Bastian, 2006; Mitsuo Kawato, 1999; Mazzoni & Krakauer, 2006;

Shadmehr & Mussa-Ivaldi, 1994; Smith & Shadmehr, 2005; Thoroughman & Shadmehr, 1999; Tseng, Diedrichsen, Krakauer, Shadmehr, & Bastian, 2007) including, as an explanation of why people are unable to tickle themselves (Blakemore, Wolpert, & Frith, 2000). The key concept behind these models is the idea of foresight or predicative processing, a concept that existed in the 19th century (James, 1890), although now there is dispute on the exact terminology and subtitles of definition (Bubic, Cramon, & Schubotz, 2010). The intrinsic link between perception and action allows for the ability to predict future actions and outcomes and modify behaviour accordingly, leading to a range of survival and performance benefits (Bar, 2007; Butz & Pezzulo, 2008; Gilbert & Wilson, 2009; Kveraga, Boshyan, & Bar, 2007; LaBerge, 1995; Llinás, 2002).

Forward models are a predictive internal model that mimics the flow of a system process by predicting its next state as the estimated sensory feedback (termed Corollary Discharge). This prediction is made given the motor command (termed efference copy) and the current state. The current state is received as sensory signals which arise from two causes, environmental effects on the operator (termed exafference) and the proprioception of the operator (termed re-afference)(Festinger & Canon, 1965; Jordan & Rumelhart, 1992; Scott Kelso, 1977; Sperry, 1950; Wolpert et al., 1995). A simple representation of a forward model is seen in Figure 2.5.

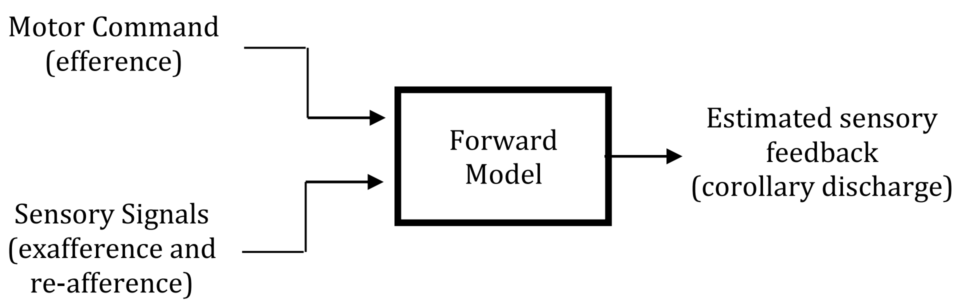


Figure 2.5: Information processing adapted from a Forward Model

Physically any motor signal from the central nervous system to the peripheral nervous system is termed an efference, likewise any sensory information signals from the peripheral nervous system to the central nervous system is termed afference (these terms are based on the terms for nerves leading into the central nervous system and nerves leading out of the central nerves system being termed afferent nerves and efferent nerves respectively).

When an efferent signal is produced and sent to the motor system, it has been suggested that a copy of the signal, known as an efference copy, is created so that exafference (sensory signals generated from external stimuli in the environment) can be distinguished from re-afference (sensory signals resulting from an operator's own actions) (Gallistel, 1981). This efference copy, providing the input to a forward internal model alongside the afference input, is then used to generate the predicted sensory feedback that estimates the sensory consequences of a motor command (Miall & Wolpert, 1996).

Complementary to forward models, inverse models attempt to estimate how to achieve a particular perceptual outcome in order to generate the appropriate motor plan. Inverse models use the desired and actual position of the body as inputs to estimate the necessary motor commands which would transform the current position into the desired one (Mitsuo Kawato, 1999; Miall & Wolpert, 1996). A simple representation of the inverse model is shown in Figure 2.6.

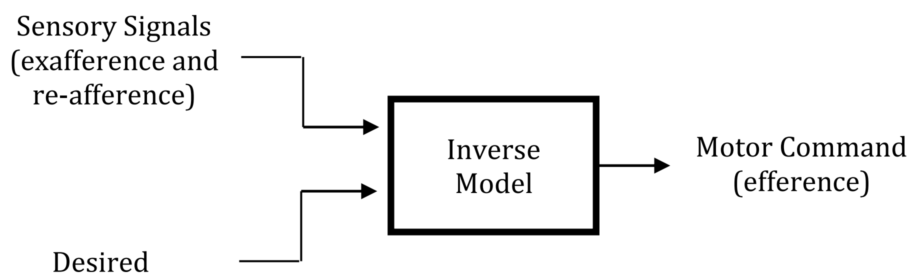


Figure 2.6: Information processing Inverse Model

The theoretical combination of inverse and forward models if they exist in the central nervous system would allow it to take a desired action and accurately control the various movements involved with said action (Mitsuo Kawato, 1999).

An example of combining the inverse and forward models for learning from error is seen in Figure 2.7 from Miall and Wolpert (1996). This combination of internal models explains that actions proceed when input of sensory information defining the current state, and input of a known desired state into the inverse model creates a motor command. This motor command is then processed by the motor system creating an action and a new achieved state. This achieved state can then be compared to the original desired state allowing any difference between the states noticed by sensory feedback to be processed by the forward model, formulating a prediction of the motor errors made during the performance of the task (Jordan & Rumelhart, 1992).

It is difficult to discern the effects of a forward model and an inverse model in a system because adaptation of either model within a system may strongly influence behaviour in early stages (Bhushan & Shadmehr, 1999).

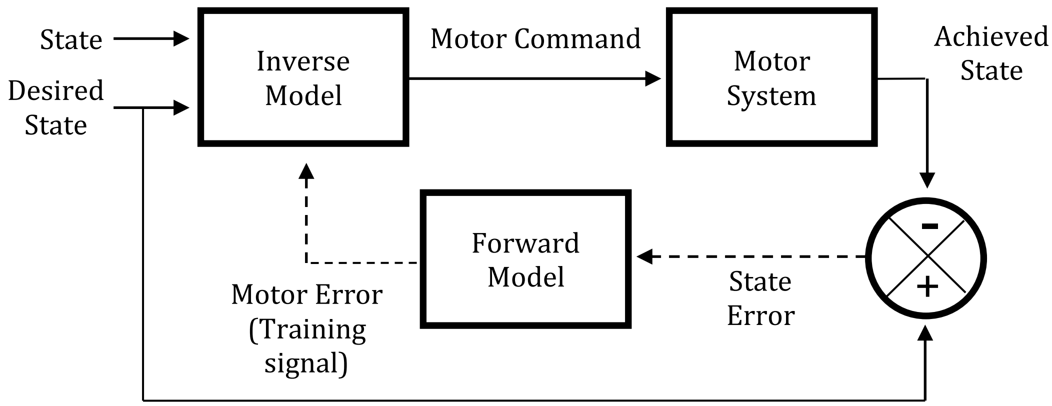


Figure 2.7: Information processing model adapted from Miall and Wolpert (1996)

The terms forward model and inverse model each represent a significant cognitive process. Comparing Figure 2.7 and Figure 2.1, Figure 2.7 essentially describes how a person is able to create a motor command based on their intentions and the information of the world around them. Although neither the forward model nor the inverse model account for memory influencing their outputs like Figure 2.4, the way Wickens & Holland's Model (Figure 2.4), does. This is a crucial flaw as previous experience has a significant impact on all human actions.

The forward and inverse models act as black boxes summarising aspects of the stimulus to response process, this may be seen as oversimplification but it also allow for emphasis on more key variables that influence the internal models.

2.6.3 Comparison of States

Through internal models it is suggested that a predicted state exists which describes the future physical position of the body or environment (Grush, 2004; Johansson & Westling, 1984; Wolpert et al., 1995). This predicted state can then be compared to the actual physical state so any differences can be identified, this allows for reassessment of the situation and provides information for a subsequent action (Bubic et al., 2010).

States most commonly represent the physical dynamics and properties of a system, for example a person moving their hand would have a state described by their joint angles, velocities, mass, skin temperature, etc. (Miall & Wolpert, 1996; Wolpert, Doya, & Kawato, 2003). A state may also refer to interaction with the environment in a non-physical way such as social interaction. Here a person's own mental state is monitored and modeled in addition to their physical state. As well as this predictions are made about their social partners mental state based on sensory feedback of their observed physical state, all of which culminates to create the appropriate communicative signals (Wolpert et al., 2003).

For the purposes of this section a state will be loosely defined as the sensory interpretation of the physical state of a person performing the action and any influences of environment on them.

Figure 2.8 from Bubic et al. (2010) depicts a simple internal model system that highlights a comparison between the predicted sensory feedback that is the predicted state, and the actual sensory feedback that is the actual state. Bubic et al. (2010) describes a 'match' or a 'mismatch' resulting from the state comparison; if a match is achieved the operator is able to continue on to their next planned action, if a mismatch occurs they can then perform an action based on sensory feedback of unexpected actual state that will compensate for this unexpected result.

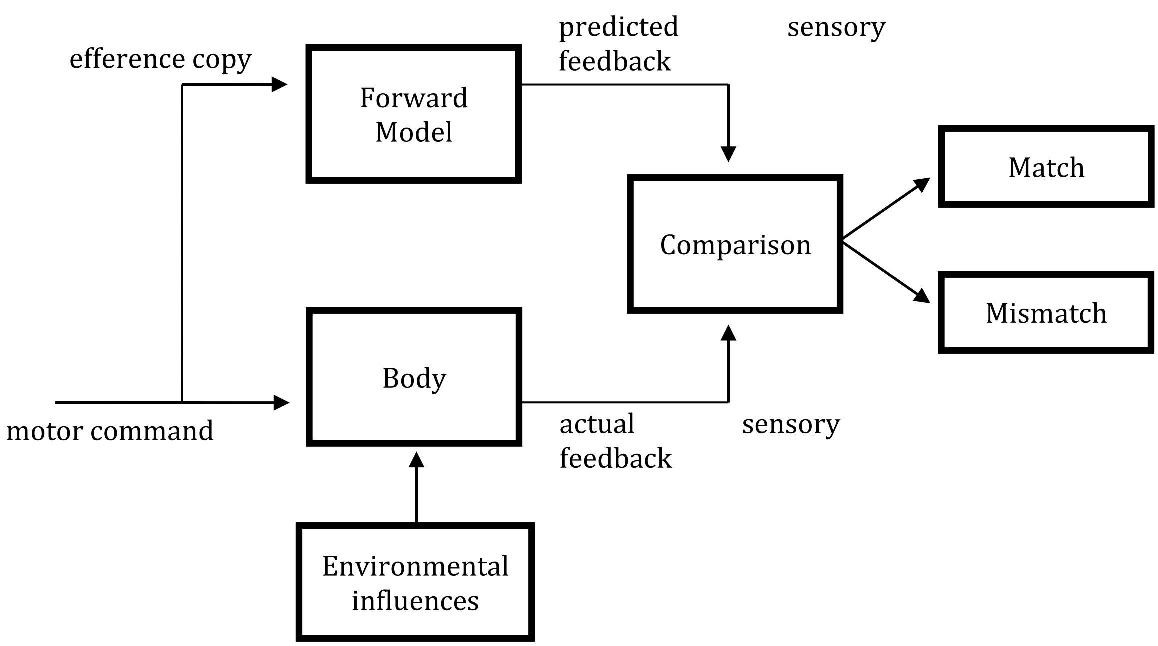


Figure 2.8: Information processing model adapted from Bubic et al. (2010)

The importance of sensory feedback to allow the mental construction of the actual state for comparison to the predicted state has significant evidence behind it. Subjects who were required to perform accuracy tasks with their hands were severely impaired when deprived of proprioceptive and cutaneous cues (Ghez, Gordon, & Ghilardi, 1995; Ghez, Gordon, Ghilardi, Christakos, & Cooper, 1990; Hasan, 1992; Miall, Haggard, & Cole, 1995; Rothwell et al., 1982; Teasdale et al., 1993). State comparison, particularly identification of discrepancies between the predicted and actual state are thought to be of significant value and are considered a main learning force (R A Rescorla, 1972; Schultz, Dayan, & Montague, 1997). Such discrepancies may identify changes in the body or environment (Grush, 2004), cause an update of knowledge and behavioural adaptation (Bubic et al., 2010; Winkler & Czigler, 1998; Winkler, Karmos, & Näätänen, 1996) and allow for more efficient use of mental resources (Corbetta & Shulman, 2002; Escera, Alho, Schröger, & Winkler, 2000).

State comparison provides a means to identify whether ones actions create the desired outcomes and allow the next action to be more effective based on the previous actions. States defined by the literature generally focus on the internal system of an operator or how the environment affects the physicality or mentality of the operator. In practice states are difficult to measure using this definition and ultimately the state of the operator exists to create or react to change in the environment around them. Because of this the majority of experiments which test internal models measure state change in the environment for example if the target ball was moved to the correct new area, rather than of the subject.

The model by Bubic et al. (2010) seen in Figure 2.8 simply and clearly identifies the comparison of the predicted and actual states which results in match or a mismatch. The model does not show feedback explicitly from the comparison to inform the next action nor does it show the motivation determining the initial motor command. These two aspects are important in understanding how an operator would complete multiple sub tasks where each action influences the next and affects predicted states to achieve their overall goal.

2.6.4 Stage and Error Models

The majority of models representing human information processing are based on the Stage Theory Model by Atkinson and Shiffrin (1968), also known as the Atkinson-Shiffrin Memory Model (Lutz & Huitt, 2003). The stage model represents the stimulus to response path, taking into account three stages of memory (Atkinson & Shiffrin, 1968). These stages of memory are described as follows with distinctions between short-term / working memory and a brief overview of the different theories regarding long- term memory

- Sensory memory represents an organism's ability to receive information from the environment via sensory receptors and store it in the central nervous system just long enough to be stored in short-term memory (Carlson et al., 2009; Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1989; Neisser, 1967).

- Short-term memory represents a system that holds a small amount of information readily available for recall for a short period of time (Atkinson & Shiffrin, 1968; Broadbent, 1958; Cowan, 2008). More commonly, working memory is used which is a combination of short-term memory and other processing mechanisms that allows an organism to hold multiple pieces of transitory information in the mind, where they can be manipulated (A. Baddeley, 2003; A. D. Baddeley, Thomson, & Buchanan, 1975; Miller, 1956).
- Long-term memory is perhaps the least understood of the stages of memory discussed here; there are several theories to describe long-term memory. The duel store memory model which ties in with the stage model describe here by Atkinson and Shiffrin (1968) suggests that rehearsals of information in short-term memory will strengthen that information's presence in long-term memory. An alternative model was made by Baddeley et al. (1975) and more recently updated by Baddeley (2000), which presents a long-term memory model based on control system which determines the flow of information know as the 'central executive'. These are linked to slave systems, which are short-term storage systems each dedicated to a specific content domain such as language or vision.

The Stage Model illustrates an external stimulus as an input into sensory memory which in turn inputs into working memory. Working memory then creates a loop in itself and with long-term memory before creating a physical motor response, seen in Figure 2.9.

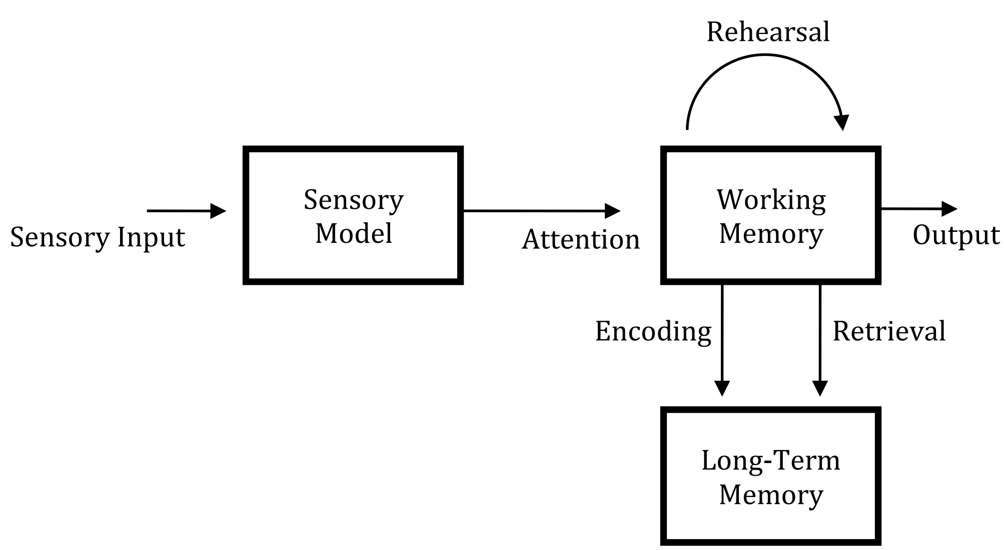


Figure 2.9: The Stage Model adapted from Gillund & Shiffrin (1984)

The Stage Model has been examined, tested, criticised and adjusted by the creators (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981; W. Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) and the academic community (R. A. Bjork and Whitten, 1974; Howard & Kahana, 1999).

The basic concept of the model has been used as a basis for explaining the concept of human information processing in simple terms in many publicly available websites and articles.

The Error Model by Helander (2005) interprets the Stage Model by Atkinson and Shiffrin (1968) combined with the three broad error types of slips, lapses and mistakes defined by Reason (1990) as seen in Figure 2.10.

The Error Model in Figure 2.9 replaces and redefines the three stages of memory from the Atkinson and Shiffrin (1968) model with stages more specific to an operator encountering and performing a task. The model also puts greater importance on output with the addition of the stage 'action execution' representing motor outputs of the operator. The stages presented in this model emulate to an extent the three performance levels identified by Rasmussen to allow interaction with the error types. Table 2.2 shows the association between the memory stages, task specific stage, the performance levels and the errors types.

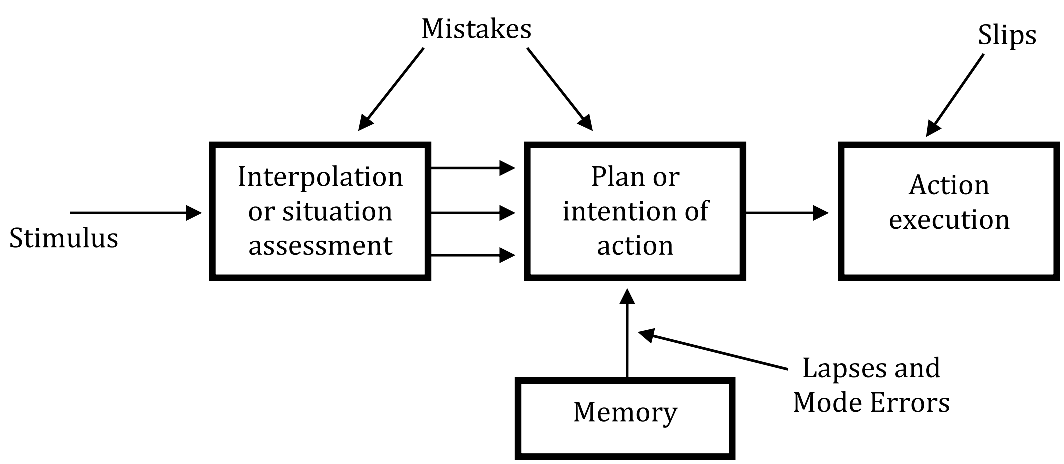


Figure 2.10: The Error Model adapted from Helander (2006) and Reason (1990)

The Stage Model and subsequently the Error Model are based on cognitive process. Although having significant empirical evidence in their favour, the Stage and Error models remain hypothetical in the sense that they are often the dependent variable being tested. As opposed to a more tangible and easily measured independent variable.

The Error Model lends itself to more practical application as identification of the various error types can then lead to potential causes for those errors. For example forgetting a step in an operation (a lapse error) is shown to interfere with the recall of memory during the planning stage in the Error Model, Figure 2.9. Therefore in this example an observed lapse error can be traced back to poor recall and a means to reduce this error at its cause can be more easily found.

Table 2.2: The associated links between a person's ability, the environment and likely error to occur.

Association between ability, environment and error			
Memory Model Stage	**Task Specific Stage**	**Performance Level**	**Linked Error type**
Sensory Memory	Interpretation or situational assessment	Rule Based	Mistake (rule based)
Working Memory	Plan or intention of action	Knowledge based	Mistake (knowledge based)
Long Term Memory	Memory	Rule Based	Lapse (including mode errors)
Output	Action Execution	Skill based	Slip

The description of these cognitive processes is simply illustrated in both the Stage Model and Error Model. This allows a general audience and those unfamiliar with cognitive science to grasp the general concepts of stimulus affecting memory leading to an action output. The Error Model arguably does this best by removing the terms used to describe the stages of memory and replaces them with simple and specifically relatable short phrases.

The Stage model includes internal feedback loops whereas the Error Model shows none, although use of the term 'plan' does allude to the idea of iterative thinking. However neither model show a feedback loop from the physical environment back into the cognitive process within the system as seen in the Information Processing Model by Miall and Wolpert (1996) (Figure 2.7) and the Information Processing Model by Bubic et al. (2010) (Figure 2.8). The lack of emphasis on environmental influence is a flaw within the Stage and Error Models.

2.7 Gaps in the Body of Knowledge

The effect of impairment and disability on task performance is an unexplored and complex area of knowledge. Impairment is a broad term that may refer to permanent, temporary and ever changing effects on a person's cognitive and physical abilities for a number of reasons.

As an example of the unique aspect of this area is someone with only a physical impairment but cognitively apt may be understand how a task is to be performed and the resulting effects of that task. However their physical inability may prevent them from being able to easily complete this task. In the current view of human error this person may be seen as making many slip errors, although a slip can be overcome easily by giving greater attention to the task (Dekker et al., 2010; Reason, 2000), this is not the case for this person as it is their physical ability not their attention focus that is failing them. This discrepancy between their physical ability and cognitive ability may then cause great frustration which in turn causes true errors which could be mistaken for a symptom of impairment.

2.7.1 Philosophical Application of Disability Models to Design

The accepted model of disability has changed significantly over time due to the many variables which affect how society views, aids and provides for those with impairments and disabilities. This in turn affects what research is done in the related fields and how technology is used for aid.

The disability models are the guides for how people with impairments are viewed by society and given appropriate provision for. The current Bio-psychosocial model recognizes that people require interaction with the other people and world around them. Because of this aid devices are beginning to incorporate links to social media, the wider internet and other digital devices such as cameras and e-readers.

A result of the current Bio-psychosocial model is the ICF which has allowed a simple means to quantify impairment. The ICF has broad application and is generally intended to provide information for local caregivers and health care professionals who will use that information in addition to their own knowledge to make recommendations for care. The ICF is not specifically intended for use in engineering usability research and development.

A greater number of more specific tools which are able to capture data that can be applied directly to product design or increase understanding need to come from the disability models. This will aid innovation and development by providing easier research, accessibility to information and greater empathy.

2.7.2 Error and Cognitive Loading Relating to Impairment

Human error is a well explored field of research but there is a lack of literature freely available to advise on practically identifying performance levels and error types (Lucas, 2001; Johnson, 1999). In addition the cause of error due physical and cognitive impairment (be it due to disability, fatigue or complex circumstance) as opposed to the users inability or cognitive loading is not well explored.

Cognitive loading is well understood, but similar to the gaps found in the knowledge of human error, the relationship between a person's impairment and cognitive loading is an area that is not well explored. Impairments are likely to create all three types of cognitive loading to varying degrees. This and the likelihood of error occurring will affect the user's ability to form metal models of the task and product.

2.7.3 Impracticalities of Information Processing Models

The information processing models examined were able to convey the concepts of how a person receives sensory input, cognitively processes this information and creates physical output. However these models were abstract and removed from direct application to design. Those attempting to use these models to improve product usability would require knowledge in psychology and neuroscience. For many industry applications gaining agreement from all people involved with the design of a product is important and any barrier to understanding reduces the effectiveness of communication. Creating a model that is easily understood and applied to product design focusing on usability will allow for practical use and tie in with other User Experience design tools such as Personas, Scenarios and User Testing.

The reviewed information processing models are not exclusive to HCI. Developing a model that incorporates the three identified aspects of HCI (Functionality of the system, Users ability and Communication) will allow the model to have greater practical application.

2.7.4 Opportunities for Advancement in the Field

There is a need to gain a better understanding of how impairments relate to human computer interaction, and for quantitative information that will have direct influence on engineering design. But it is likely that a means to allow engineers to understand the mindset, approach to a task and the process taken to understand and perform that task is required first and may be more valuable. This may be achieved through the knowledge gaps identified here; the need for practical application of disability models, reducing the barriers to understanding and applying information processing models and a better understanding of how impairments relate to error, cognitive loading and the formation of metal models.

This need for a better understanding of how a person's impairment influences task performance gave rise to the first three purposes of the study: Identify the effects of impairments on a user's measured performance, self-reported performance and usability experience; determine how cognitive loading affects a user's measured and reported performance and identify what types of errors are made when using the controller and the possible cause for these.

The final purpose of this study is to create a model that can be used to to describe the process of completing a task which takes into account the external and internal factors affecting that process. This model will be based on the reviewed literature as well as the results of this study.

3.0 Methodology

The practical experiments of this study are set out with descriptions and reasoning for the experimental design including the physical test apparatus, how the controller is operated, what tasks were done on the controller and gaining ethics approval.

Then the types of participants sought and their recruitment are addressed. The various research tools are examined including observational methods, assessments given to the participants and statistical analysis methods. Finally the full experimental procedure is summarised

3.1 Research Questions

Based on the identified knowledge gaps, the three determined aspects of HCI and purposes of this study, the following research questions were put forward to direct the experimental design and analysis.

1. Which impairment type most affects HCI task performance?
2. Which error types are most prevalent in HCI when comparing type and severity of impairments?
3. Is there a difference in participant reported usability compared to participant reported task satisfaction and measured task performance?

3.2 Approach

To best answer the research questions the appropriate instruments for measurement and data collection were selected while designing the experiment. Following this the appropriate statistical analysis tools were chosen to extract information from the data in the most effective and direct way. A number of methods for analysis of controller and measurement instruments were reviewed; see Appendix 1.

The chosen instruments for gathering data were the Task Load Index (TLX), the International Classification of Functioning (ICF), System Usability Scale (SUS), various demographics, and task performance times. These and the statistical analysis methods used are explained further in section 3.6.0.

Constraints on the experiment were predominantly based around the practical considerations of the hardware and participants. Additionally the time taken to complete the experiment, creating a portable rig and the available resources at the university limited the study.

Within these constraints specific features of the controller were chosen to be used as tasks to best capture the data, fulfilling the study purpose and answering the research questions.

3.3 Measured Variables

The experiment was designed to capture a number of different variables based on the measurement tools discussed above in section 3.2.0. These variables were chosen with the intention to answer the above research questions and fulfil the study purposes.

Table 3.1 the measured variables are presented and categorised based on the measurement tools. The variables were not all used in the analysis and some variables were condensed by combining the data for the analysis. See Appendix 13 and 14 for the full raw data and condensed raw data and Appendix 8, 9 and 10 for the questionnaires relating to these variables. In total there are 61 raw measured variables for each of the 40 participants.

Table 3.1: Summary of measured variables

Demographic Questions	ICF levels	Performance Times*	Task Load Index (TLX)**	System Usability Scale***
Age	Quality of psychomotor functions	Task 1 Chair, Attempt 1,2,3, and 4	Mental Demand	I think that I would like to use this system frequently
Gender	Control of simple voluntary movements	Task 2 Lights, Attempt 1,2,3, and 4	Physical Demand	I found the system unnecessarily complex
Experience using powered wheelchairs	Auditory perception	Task 3 Clock, Attempt 1,2,3, and 4	Temporal Demand	
Time spent using powered wheelchairs	Visual perception	Task 4 Cog, Attempt 1,2,3, and 4	Self-reported performance	
Confidence with technology	Tactile perception		Effort	
	Acquiring skills		Frustration	

*Time taken to complete each attempt of each task

**Given after attempt 1 and attempt 4 of each task

***Given to participants at the end of the experiment

3.4 Design of Experimental Hardware

This design was based on preliminary tests which were used to determine an appropriate number of tasks to test, roughly how long the experiment would take and to practice the use of the various assessments.

3.4.1 Test Rig

A rig was built using the wheelchair controller separated from a wheelchair. This was done for practical purposes allowing the rig to be portable. This may have reduced the realism of the experience but its practicality was required.

Figure 3.1: Labeled experimental wheelchair controller rig.

Testing was done at various locations around Christchurch and many participants were in wheelchairs. Transporting a powered wheelchair and moving some participants from their chair into the test chair would have complicated the experimental process. It would have created an unpleasant experience for participants and increased the hazards and risks of the experiment. In addition the controller on a wheelchair would not be easily reconfigurable for left handed participants compared to the test rig.

The Test rig used can be seen in Figure 3.1. The rig consisted of the controller which sat on a custom built adjustable mount that included a pre-made tripod. A webcam connected to the mount recorded the users operating the controller. The webcam connected to a laptop while the controller connected to a control box supplied by the industry partner which mimicked the motors and other hardware of a wheelchair.

3.4.2 Controller Operation

The controller is operated through ten push buttons as seen in Figure 3.2. The six black paired buttons are used to operate the Graphical User Interface (GUI), the two on the left hand side labelled + and - are used for the speed controls. Of interest are the other four black buttons, the middle horizontal pair labelled I and II, and the right vertical marked < and >, these operate the menus containing the non speed and driving related features. In addition to these buttons the joystick is used to activate the feature shown in the centre of the screen.

The participants only operated the non drive features and always started on the same screen for consistency, the starting screen can be seen in Figures 3.3 to 3.6 as the first screenshot in each image.

3.4.3 Controller Tasks

In this study only the additional features of the controller were being assessed rather than the driving and handling controls. The controller has over a dozen features depending on how the controller is set-up features. Only some of the features could be tested to keep the experiment time practical.

Figure 3.2: Labelled controls. Not shown is the joystick which is required to select the central icon on screen.

The button controls are shown and labelled in Figure 3.2. The joystick (not shown in Figure 3.2) was only needed to select the central icon by pushing it forward.

Four tasks were selected to be performed by participants as seen in Table 3.1. The first three tasks were features of controller and the fourth task consisted of the three previous tasks with a word association test to measure performance under cognitive loading.

Table 3.2: Summary of Tasks

Task	Reference name	Goal of task	Minimum Actions Required to complete
1	Task 1 Chair	Raise the chair	4
2	Task 2 Lights	Turn the lights on	5
3	Task 3 Clock	Turn the clock off and then on	8
4	Task 4 Cog	Complete tasks 1 through 3 with word association test	15

Participants were given four attempts at each task done consecutively, beginning with four attempts at Task 1 Chair before moving on to attempt Task 2 Lights four times, and so one. This made a total of 16 attempts over all tasks.

Task 1: Raise the Chair

This task was chosen as it is one of the most fundamental features of any powered wheelchair and ideally should be one of the easiest and fastest to activate. The chair may be adjusted in multiple ways, tilted, raised and leaned. Testing to see if participants could distinguish the rise function from the others was of interest to the industry partner.

In practice the experiment participants did not have to 'hold down' the raise chair feature but briefly bring it to the active screen and then release (if there was a chair it would slowly raise). Figure 3.3 shows screen shots and actions to complete Task 1 Chair. Participants knew the task was completed when the chair icon turned green as seen in the final screen shot of Figure 3.3.

Figure 3.3: Stages of Task 1 Chair with required next actions indicated in yellow

Task 2: Lights On

Activating the lights is another important feature for many wheelchairs. Unlike raising the chair it is unlikely to be used often but when used it will be of importance. It is likely that the person will likely be in an outdoor and/or late night environment when the user is short on time.

Figure 3.4 shows screen shots and actions to complete Task 2 Lights. Participants knew the task was completed when a small green circle with a light icon appeared at the top of the screen as in Figure 3.4.

Figure 3.4: Stages of Task 2 Light with required actions indicated in yellow

Task 3: Turning the Clock Off and On

This was chosen as it is known to be difficult by the industry partner. As can be seen in a number of the screen shots in Figure 3.5 there are several ticks and crosses on the screen making it unclear which would turn the clock 'on' or 'off'.

The task began with the clock 'on' for participants to identify where it was on the screen and so they would know when it was 'off'. Participants were asked to turn the clock back 'on' after turning it 'off' to make the task more difficult as well as to reset the clock to 'on' for the next attempt.

Although this task is not as critical as the other two features it is arguably one of the most used as the clock may be glanced at several times a day. The clock may need to be changed at least twice a year for daylight saving making this a unique feature as repetition of the task would likely not be regular enough to memorize the process.

Figure 3.5 shows screen shots and actions to complete Task 3 Clock. Participants knew the task was completed when the digital clock seen in the top right hand corner of the screen disappeared and that it was back on when it reappeared.

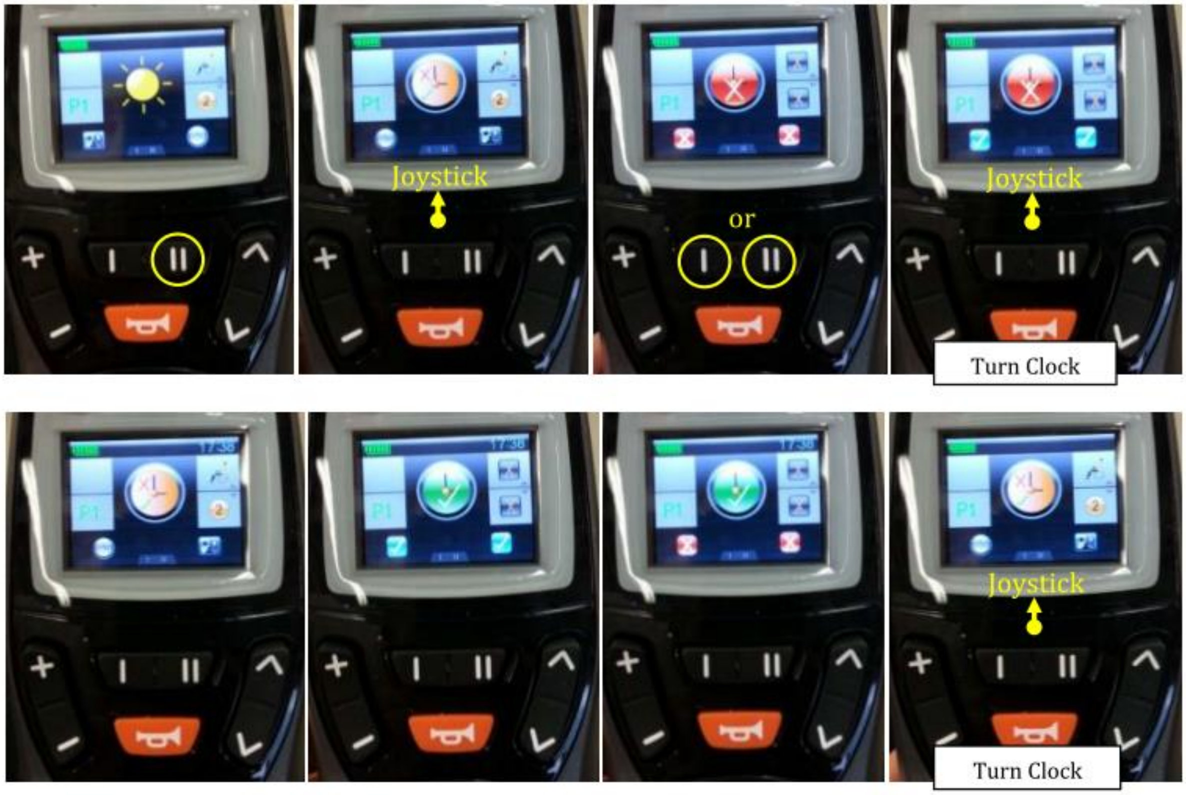


Figure 3.5: Stages of Task 3 Clock with required next actions indicated in yellow

Task 4: All Tasks with Cognitive Loading

The fourth and final task required participants to repeat all previous tasks (raise the chair, turn the lights on and turn the clock off and on) with cognitive loading applied in the form of a word association test.

The word association test involved the researcher saying a single word and the participants replying with a different single word as quickly as possible with the first word that came to mind. For example if the researcher was to say 'Dog' the participant may say 'Cat' or 'Bark'. The words used can be seen in Appendix 3, the words chosen were selected from (Church & Hanks, 1990) where they were categorised as neutral.

The words were read out as soon as the participant had replied to the previous word; this meant that the participants who answered faster would in total be given more words. This method was used to ensure participants were always engaged with the intent of keeping a more consistent cognitive load. On some occasions participants were asked the whole word list, if this was the case words from the list were picked at random.

The previous tasks were well practised meaning there would be less of an interaction between the cognitive load to learn a new task and the cognitive load of the word association test. All the previous tasks were used to ensure that the task would be a substantial length allowing for the effects of the cognitive loading to take effect.

Between each of the three tasks within task four the beginning screen was not reset, this meant that participants were less restricted on how they moved from task to task, potentially adding to the cognitive loading in a less controlled manner compared to the word association test. The most efficient method of completing Task 4 Cog is shown in Figure 3.6.

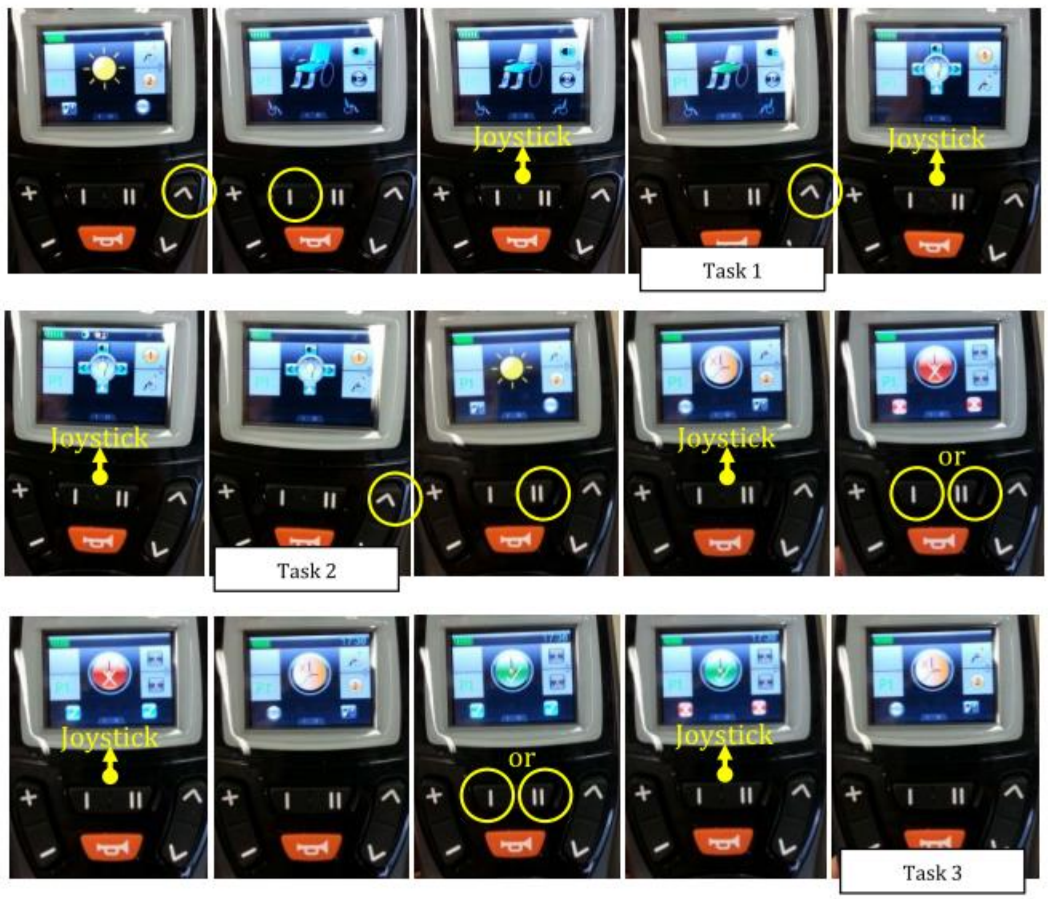


Figure 3.6: Stages of Task 4 Cog with required next actions indicated in yellow

3.4.4 Ethics Approval

Ethics approval was gained from the University of Canterbury Human Ethics Committee.

As this study required participants who have impairments a high risk application was made to the University of Canterbury Human Ethics Committee. The approval and application forms for the study and the participant information sheet and consent form can be seen in Appendix 4.

An application was attempted for the New Zealand Health and Disability Ethics Committee (HDEC) but their approval was not needed as the study was of low enough risk. See Appendix 5 for the requirements needing HDEC approval compared to the requirements of this project.

Participants were reward for their time with a chocolate bar and $10 shopping mall voucher.

3.5 Participants

Here the key variables are identified that determined what type of participants would be sought. The division of the sourced participants into the three groups for analysis purposes is then explained.

3.5.1 Key Variables

The primary purpose of this study is to determine the affect of impairments on usability. In addition age was thought likely to have a strong effect on usability and needed this to be taken into account.

Participants differing in the two key variables of age and impairment were sought out. As this study had limited on resources it was decided that comparing extremes of these two variables was the best approach in order to obtain sufficient significant results.

This created a need to find four distinct participant groups, as seen in Figure 3.7. These four groups were young with impairment, old with impairment, young without impairment and old without impairment. A total of 40 participants were found, with a near equal spread between all groups apart from young with impairment where no participants were sourced.

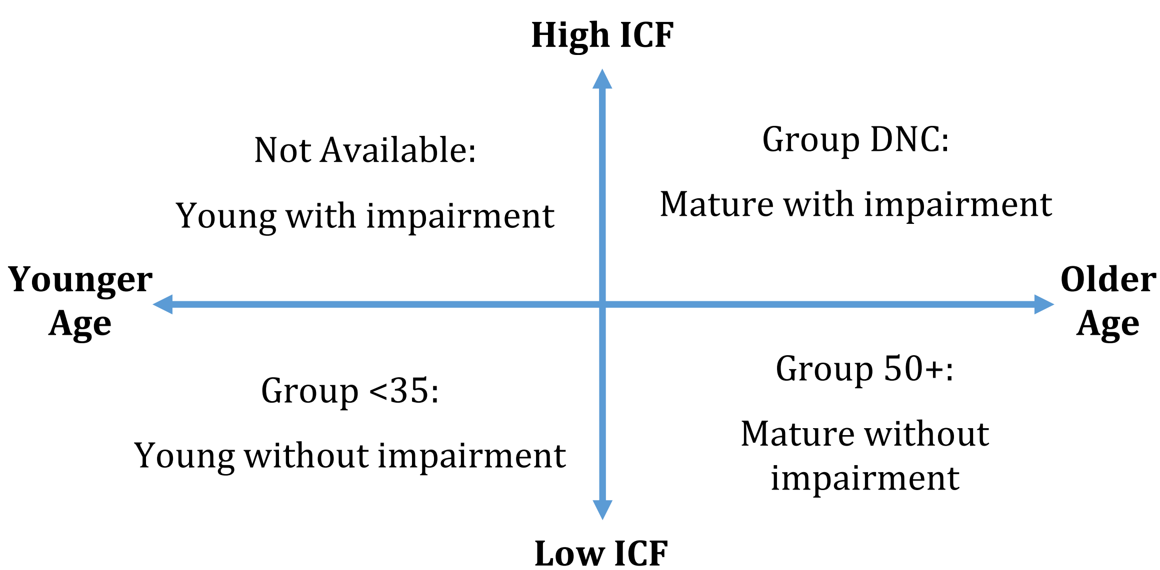


Figure 3.3: Sought participant groups

3.5.2 Sourcing and Grouping Participants

There were 40 participants in the study who fell into three clear groups. These groups matched closely the sought variables seen in Figure 3.6 above where they are labelled DNC, <35 and 50+ . The groups are participants under 35 years (<35), participants over 50 years (50+) and participants who did not complete the experiment (DNC). It is noted that participants in the DNC group generally had greater levels of impairment. Further descriptions of each group and how they were sourced follow, descriptive statistics of each group can be found in Results Section 4.2

Impairment was not used to define any group as there was a range of impairment levels compared to age which was either under 35 years or over 50+. DNC was chosen as a group to emphasise that nearly a third of all participants stopped the experiment before completion and that several statistical analysis methods were carried out solely on these participants. Compared to the other groups DNC did have an average higher level of impairment.

Group <35: Participants under 35 years

The young group of participants was sourced through the University of Canterbury. Young able bodied students were found easily using poster advertising and word of mouth. Young disabled students were sought through the University of Canterbury Disabled Resource Services[2]. Unfortunately no young people with impairments participated, despite the help of the Disability Services.

This group was termed <35 as it consisted of people under the age of 35 years. There were 15 people in this group.

Group 50+: Participants over 50 years

Mature participants without impairments were sourced predominantly through word of mouth, asking for volunteers through people who knew the researcher. Mature participants with impairments were sourced through the disabled community.

Several care centres, disability service providers and community support organisations were contacted to find participants with impairments, see Appendix 7. The participants sourced through these contacts were older.

[2] http://www.canterbury.ac.nz/disability/

All participants who were able to complete all four tasks and not in the <35 group were 50 years or older, forming the 50+ group. The 50+ group totalled 12 participants.

Group DNC: Participants who did not complete all four tasks

Participants who did not complete all four tasks were grouped together for analysis; incidentally the majority of participants in this group were older and had greater levels of impairment than the other two groups.

This group was specifically identified as 13 participants; nearly a third of the total 40 fit these criteria. This is arguably the most valuable group in the study and of most interest to the industry partner.

3.6 Research Tools

The research tools used are discussed in this section. They are divided into three types, observational methods used during the experiment, assessments given to the participants during the experiment and statistical analysis methods used for data analysis the data post experiment.

3.6.1 Observation Methods

Three different observational methods were used to collect data: recording performance times, an attempt at recording the actions of the participants and recording if assistance was given.

Recorded Performance Time

Each participant's attempt for each task was video recorded, resulting in 640 recorded videos. These videos were then used to determine the time taken for each attempt. This was done instead of recording the time directly during the experiment. Observing the videos allowed the start and finish times of the attempts to be more accurately identified. The researcher used their judgement to identify the start and finish of each task which was to an extent subjective.

Recording Actions (attempted but not done)

There was an attempt made to record controller inputs; the order of which buttons the participants pressed and when the joystick was activated. This would have allowed a detailed view of the participant's actions. The intention was to allow errors to be more easily identified and a comparison made between participants and the designers' expected actions.

Unfortunately recording the actions could not be done effectively, despite three different approaches being taken. Firstly the industry partner was approached about modifying the digital output of the controller to signal when the buttons were pressed to third party software. This was considered too resource demanding and impractical for the timeframe.

Secondly the internal hardware of the controller was accessed by university technicians in the hope of capturing the direct electrical signals from buttons. Unfortunately this was not possible.

Finally pressure pads were externally attached to the buttons of the controller. This was trialed in preliminary experiments but proved unreliable and appeared to affect how the participants used the controller.

Recording Assistance

During the experiments the researcher recorded whether assistance was given to the participants. What constituted 'assistance' was subjective to the researcher at the time of the particular experiment. This measure was recorded as a rough indicator for the groups that were assisted and these may have a greater performance recorded than they would otherwise.

3.6.2 Participant Assessments

The assessments given to the participants were primarily handed to them on paper with the participants reading the questions themselves and writing the answers. On some occasions the assessments were read to participants if they were unable to use a pen or read easily. On a few occasions the assessments were filled in by a nurse. Participants were free to ask any questions about the assessments.

The following assessment tools were used and their application in this study is described.

Demographics

The first questions asked of patients were of a general nature not associated with an established assessment. These questions asked their:

- Age
- Gender
- Experience using an electric wheelchair
- Confidence with using new technology

For a copy of the questions given to the participant before beginning the tasks see Appendix 8.

The participant's exact age was asked. Options for gender were either male or female. There was a near even spread with 19 female and 21 male participants.

Experience using a wheelchair was made up of two questions with several options each. The participants were asked the extent of experience they had and how long ago that experience occurred. For analysis the results of these questions were condensed to be either 'had experience' or 'did not have experience', this was done at the researchers' discretion due to the split results among the participants. Participants who had more than one continuous day of experience using a wheelchair were considered as having experience; a total of 12 participants met this criteria with most of these participants in the DNC group. The remaining 38 participants had very little to no experience using an electric wheelchair and were considered to have no experience.

Confidence with using technology was gauged by asking participants to indicate their confidence on an 11 point scale (0 to 10).

ICF - International Classification of Functioning, Disability and Health

The International Classification of Functioning, Disability and Health (ICF) was used to qualitatively measure levels of impairment with reference to operating a wheelchair controller. The World Health Organisation first published the ICF in 1980 and endorsed it in 2001 as a multi-purpose classification intended for a wide range of uses in different sectors; it is currently in its tenth revision ("WHO | International Classification of Functioning, Disability and Health (ICF)," 2014).

The ICF uses a 6 point scale to measure intensity of an individual impairment, 2 of the points were not used in this study as these related to impairments where severity cannot be specified and inapplicable impairments. Table 3.3 shows the 5 point scale used in the ICF with definitions as presented to participants.

The framework of the ICF is divided into four sections; body functions, body structures, activities and participation and environmental factors. Each of these sections is further divided into chapters, and further sub-sections, identifying hundreds of functions[3] (Rentsch et al., 2003). Note that the term functions describes an ability of a person, the term impairment is used to identify a person's function that scores greater than zero on the scale seen in Table 3.2.

[3] http://apps.who.int/classifications/icfbrowser/

Table 3.3: ICF measures

ICF extent of impairment measure	
0	No impairment, you have no difficulty
1	Mild impairment, a difficulty is present less than 25% of the time, with an intensity you can tolerate and which has happened rarely over the last 30 days.
2	Moderate impairment, a difficulty that is present less than 50% of the time, with an intensity, which is interfering in your day to day life and which has happened occasionally over the last 30 days.
3	Severe impairment, a difficulty is present more than 50% of the time, with intensity, which is partially disrupting to your day to day life and which has happened frequently over the last 30 days.
4	Complete impairment, a difficulty that is present more than 95% of the time, with an intensity, which is totally disrupting to your day to day life and which happens every day over the last 30 day

Six functions were identified which would affect a person's ability to operate the controller and any electronic device in general; these are seen in Table 3.4. The six functions were chosen based on the aspects of human computer interaction examined in the literature review section 2.2.0.

See Appendix 11 for the relevant sections of the ICF index referring to the six functions and the website address to the WHO ICF browser.

Table 3.4: Chosen ICF measures

Function	**ICF index number**	**ICF Description**	**Question given to participants**
Quality of psychomotor functions	B1471	Mental functions that produce nonverbal behavior in the proper sequence and character of its subcomponents, such as hand and eye coordination, or gait.	At what level are your motor functions in performing sequenced movements such as typing?

Control of simple voluntary movements	B7600	Control of simple voluntary movements - Functions associated with control over and coordination of simple or isolated voluntary movements	At what level are your functions associated with control over and coordination of simple or isolated voluntary movement such as reaching to grab a glass?
Auditory perception	B1560	Mental functions involved in discriminating sounds, tones, pitches and other acoustic stimuli.	At what level are your abilities at discriminating sounds, tones, pitches and other noises?
Visual perception	B1561	Mental functions involved in discriminating shape, size, color and other ocular stimuli.	At what level are your abilities at discriminating shape, size and colour of objects and images?
Tactile perception	B1564	Mental functions involved in distinguishing differences in texture, such as rough or smooth stimuli, detected by touch.	At what level are your abilities at feeling differences in texture by touch, such as rough or smooth?
Acquiring skills	D155	Developing basic and complex competencies in integrated sets of actions or tasks so as to initiate and follow through with the acquisition of a skill, such as manipulating tools or playing games like chess	At what level are your abilities to learn and develop basic sets of actions to perform a new skill or task, such as manipulating tools or playing games like chess?

The six functions are intended to represent how a person interacts with most systems or devices. The most crucial and hardest to capture is the cognitive element of interaction; this is represented by the functions 'Acquiring skills' and 'Quality of psychomotor functions'. The others are visual, auditory and tactile interactions.

In this study the ICF assessment was given to participants at the beginning of the experiment after the demographic questions. The assessment had the 'Question Given to Participants' as well as the 'ICF Description' both seen in Table 3.3.

The scale was explained and participants were asked to answer the questions with reference to using the controller. For example when answering the Control of Voluntary Movements question if a participant had little or no control over their legs compared to their arms and hands they would answer in relation only to their arms and hands.

NASA TLX – Task Load Index

The NASA Task Load Index (TLX) is a subjective, multidimensional assessment tool that rates the perceived workload in order to assess a task, system, or a team's effectiveness or other aspects of performance. It was developed by the Human Performance Group at NASA's Ames Research Center (Hart, 2006; Hart & Staveland, 1988). The TLX uses a 20 point scale to specifically measure:

- Mental Demand: How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?
- Physical Demand: How much physical activity was required? Was the task easy or demanding, slack or strenuous?
- Temporal Demand: How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?
- Overall Performance: How successful were you in performing the task? How satisfied were you with your performance?
- Frustration Level: How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?
- Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

The TLX was given to participants after the first and fourth (final) attempt for each task. This was done to gauge the change in perceived performance and demand. In preliminary trials giving the TLX more than twice per task proved impractical and ill received by the participants.

For the purpose of analysis the TLX was broken into a performance measure and a demand measure. The Performance measure was only the Overall Performance question and the Demand measure was an average of the other variables. This was done to reduce the time of the analysis as the TLX was not a primary assessment in line with the study's purpose.

The terms used for the performance and demand measures through the rest of this report are as follow:

- TLX End Demand

  The self-reported demand of a participant from the NASA TLX assessment upon completion of the four attempts as a single task. This value is an average of mental demand, physical demand, temporal demand, frustration and effort.

- TLX Start Demand

  As with TLX End Demand but upon completion of the first attempt of a single task.

- TLX End Performance

  The self-reported performance of a participant from the NASA TLX assessment upon completion of the four attempts as a single task.

- TLX Start Performance

  As with TLX End Performance but upon completion of the first attempt of a single task.

For a copy of the TLX assessment given to participants see Appendix 9.

SUS – System Usability Scale

The System Usability Scale (SUS) is intended as a quick and dirty method for a broad measure of subjective usability from participants (Brooke, 2013). The system itself consists of ten questions, each using a ten point Likert scale (Strongly Agree to Strongly Disagree) measuring three key aspects of usability identified by the ISO standard 9241-11:1998 [4]:

- Effectiveness - can users successfully achieve their objectives
- Efficiency - how much effort and resource is expended in achieving those objectives
- Satisfaction - was the experience satisfactory

The ten questions given to participants are as follows, starting with the theme of the question which was not given:

1. Liking of System - “I think that I would like to use this system frequently”
2. System Complexity - “I found the system unnecessarily complex”
3. Ease of Use - “I thought the system was easy to use”
4. Support for Use - “I think that I would need the support of a technical person to be able to use this system”
5. Function Integration - “I found the various functions in this system were well integrated”
6. Inconsistency - “I thought there was too much inconsistency in this system”
7. Speed to Learn - “I would imagine that most people would learn to use this system very quickly”
8. Cumbersome - “I found the system very cumbersome to use”
9. Confidence in use - “I felt very confident using the System”
10. Amount to Learn - “I needed to learn a lot of things before I could get going with this system”

The SUS was given to participants after all tasks had been completed but before the close out interview. In this study the SUS was used mostly in analysis as an averaged score for all ten questions with consideration of positive and negative framing (SUS Average) or less frequently by each question.

There were some limitations of the SUS. Based on observation of the researcher not all participants understood the context or meaning of the questions with regard to the controller. Participants were unsure of what the questions were asking exactly, particularly SUS Function Integration and SUS Inconsistency. Participants also appeared to answer the SUS questions with reference to the general population rather than themselves. For example one participant who

[4] https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en

was in the DNC group and gave a SUS Average score of 5/10 said "I couldn't really use it but I think a young person who is used to computers would have no problem."

For a copy of the SUS assessment given to participants see Appendix 10.

Close Out Interview

The final assessment was the close out interview, the purpose of this was to allow participants to give their own feedback and opinions about the controller and to allow the participants to self-identify errors they made. Several questions were asked in the following order:

- General Comments

Asking for feedback, what the participant liked, did not like and what they would like to see added. This was intended to promote discussion.

- Physical Layout

Asking for the participants' thoughts and opinions on the controllers physical layout, the button joystick, screen size etc

- Icon Clarity

Asking for the participants' thoughts and opinions on the clarity of the menus icons i.e. easy to see and understand.

- Mistakes

Were there any times when the participant was very confused and thought they made any wrong choices and on which tasks and attempts did this occur. This was to help determine the frequency of mistakes as defined by Reason (2000).

- Lapses

Were there any times when the participant had trouble recalling what they did to complete the task on a previous attempt, On which tasks and attempts did this occur. This was to help determine the frequency of lapses as defined by Reason (2000).

- Slips

Were there any times when the participant had trouble hitting a button or by accident pressed a button they weren't intending to hit, on which tasks and attempts did this occur. This was to help determine the frequency of slips as defined by Reason (2000).

For a copy of the close out interview questions given to participants see Appendix 10.

3.6.3 Statistical Analysis Methods

A variety of statistical tests were used to analyse the data. These are briefly explained and their use in this study outlined.

Independent T-test

The Independent T-test determines whether there is a statistically significant difference between the means in two unrelated groups and allows the effect size of the difference between means to be calculated.

In this study an Independent T–test was used for two different statistical analyses. The first was to compare all measured variables between the three identified groups.

The second was to compare ICF Specifics between those participants who completed all the tasks and those who did not (groups <35 and 50+ combined compared to DNC). It is noted that the ICF measures are not fully independent, therefore weakening the analysis. However useful information can still be gained particularly in conjunction with other the analyses done in this study.

Stepwise Regression

Stepwise regression includes regression models in which the choice of predictive variables is carried out by an automatic procedure; this creates a model that includes only the strongest independent variables.

In this study stepwise regression was used to identify the measured variables that most contributed most to the SUS scores. Multiple regressions (explained below) was used alongside Stepwise regression to get a better understand of what contributed to the SUS results.

Multiple Regression

Multiple regression is used to predict the value of a dependent variable based on the value of two or more independent variables. It determines the overall fit of a model and the relative contribution of each of the predictors to the total variance explained in the dependent variable.

Multiple regression was used in this study to predict three dependent variables. Firstly TLX End Demand and secondly TLX End Performance for each of the four tasks based on the Performance Time of each attempt. Thirdly SUS predicted separately by the TLX Start Demand for each task, TLX End Demand, TLX Start Performance, TLX End Performance, Performance Time for Attempt 1 of each task and Performance Time for Attempt 4 of each task.

The method analysis used in this study required adding independent variables to the model in chronological order resulting in a total of 96 regression analyses. Family wise error was considered but was not deemed to be a problem as the analyses were not all related and a number of them were excluded due to not meeting the necessary assumptions for multiple regression.

Table 3.5: Experimental variables

Dependent Variable		**Independent variables (In chronological order as added)**			
TLX End Demand*		Attempt 1	Attempt 2	Attempt 3	Attempt 4
TLX End Demand*	TLX Start Demand	Attempt 1	Attempt 2	Attempt 3	Attempt 4
TLX End Performance*		Attempt 1	Attempt 2	Attempt 3	Attempt 4
TLX End Performance*	TLX Start Performance	Attempt 1	Attempt 2	Attempt 3	Attempt 4
SUS		TLX Start D Task 1	TLX Start D Task 2	TLX Start D Task 3	TLX Start D Task 4
SUS		TLX End D Task 1	TLX End D Task 2	TLX End D Task 3	TLX End D Task 4
SUS		TLX Start P Task 1	TLX Start P Task 2	TLX Start P Task 3	TLX Start P Task 4
SUS		TLX End P Task 1	TLX End P Task 2	TLX End P Task 3	TLX End P Task 4
SUS		TLX Start P Task 1	TLX Start P Task 2	TLX Start P Task 3	TLX Start P Task 4
SUS		TLX End P Task 1	TLX End P Task 2	TLX End P Task 3	TLX End P Task 4
SUS		Attempt 1 Task 1	Attempt 1 Task 2	Attempt 1 Task 3	Attempt 1 Task 4
SUS		Attempt 4 Task 1	Attempt 4 Task 2	Attempt 4 Task 3	Attempt 4 Task 4

*repeated for each Task

An example of application is to predict the TLX End Demand for Task 1 Chair using the attempts for Task 1. The analysis began with only Attempt 1, and then repeated using Attempt 2 and so on until Attempt 4.

The dependent and independent variables for the multiple regressions are seen in Table 3.5.

Linear Mixed Effects Model

Mixed models account for both fixed and random effects predicting a dependent variable in the same analysis. The model allows differences within and between the fixed effects to be seen. It is noted that this analysis is done using stacked data (otherwise know as long form).

In this study the dependent variable was Performance Time with the three Groups, four Tasks and four Attempts as fixed effects. The participant number was the only random effect.

The analysis was used to determine which of the fixed effects significantly affected Performance Time.

Cox Regression Survival Analysis

Cox Regression creates a survival function for 'time to event' data that predicts the probability that a predetermined event occurs at a given time based on set predictor variables.

For the purposes of this study the 'event' was a participant stopping the experiment, either the participant asking to withdraw or the researcher stopping the experiment due to the participant's inability to complete the task. This second situation arose only in the most extreme circumstances, for example when a participant was taking longer than 45 minutes to complete a task or clearly did not know how to operate the basics of the controller.

The predictor variables were Age and ICF. Several Cox Regression analyses were performed, firstly to determine impact of Age compared to ICF, then to determine the impact of the ICF Specifics.

Content Analysis (Qualitative analysis)

Content analysis is used to classify qualitative information into themes or categories from empirical documentation, where this documentation is often written text.

In this study a simple content analysis was done on the notes recoded from the close out interview. This involved paraphrasing the various answers as single word themes and then counting these words. This analysis on its own is relatively weak within this study but is used as a comparison to the various quantitative methods.

3.7 Experimental Procedure

For each of the forty participants the experimental procedure was kept as consistent as possible, an outline of the procedure can be seen in Table 3.5, for the full procedure see Appendix 12.

Table 3.5: Basic experimental procedure

1. Welcome and gain ethics consent
2. Give Demographics and ICF assessments
3. Explain basic operations of controller
4. Task 1
    a. Explain objective of Task
    b. Turn on controller and bring to start screen
    c. Ask participant to complete Attempt 1
       (turn off controller upon completion)
    d. Explain and give TLX assessment
    e. Reset to start screen and ask participant to complete Attempt 2
    (turn off controller upon completion)
    f. Repeat step e. until completion of Attempt 4
    g. Give TLX assessment
5. Repeat step 4. Until completion of Task 4
6. Explain and give SUS assessment
7. Explain and give Close Out Interview
8. Thank and reward participants

Participants were tested from October to December 2013. They were not tested in any planned order but generally similar participants were tested adjacent to each other. This was due to the rig being used most efficiently when set up in a location for several days, as similar participants were generally at a location (such as the younger demographic being at the university) so those participants were tested over that time frame.

The length of the experiment varied greatly between participants, which was unexpected. The range of total time for the experiment ranged from about 30 minutes up to 2 or more hours.

Basic operation of the controller

Before starting the experiment a brief overview was given of how the controller was operated. This explanation was kept deliberately brief and did not involve a demonstration of the controller. This was to ensure that participants did not begin the experiment primed. Participants were told they did not need to use the left hand side paired black buttons during the experiment. A rough transcript of the overview follows:

> "The right pair of black buttons move the menu of options on the right of screen (indicate buttons).
>
> The middle pair of black button moved the menu of options seen at the bottom of the screen (indicate buttons). This bottom menu is an expansion of the right menu.
>
> The joystick when pushed forward activates the icon seen in the centre of the screen (mimic pushing joystick forward)."
>
> You won't need to use any other buttons, including the plus and minus button on the right hand side (indicate buttons)'

In addition concise explanations of the task objectives and assessments were given to the participants.

Tasks

Before each task it was explained to participants that 'on screen' feedback would be given so they would know that a task was completed. For Task 1 Chair the chair icon changed to green, for Task 2 Light a green 'light' icon appeared, for Task 3 Clock the clock on the screen disappeared/reappeared, and for Task 4 Cog the other feedback prompts all apply. For a more detailed explanation of these feedback prompts see section 3.4.3 Participant Tasks.

To reset the controller it was simply turned off and on using the green power button at the top of the controller. Turing it back on brought up the speed control menu which easily changed to the starting menu with a single button press. The starting screen is the first screen shot seen in Figures 3.3 to 3.6 in section 3.4.3.

4.0 Results

The results are presented first with reference to the purposes of the study as in the Introduction (Section 1.0) and secondly with reference to each of the statistical analysis methods used.

This dual reporting was done as multiple statistical methods relate to each proposal and some methods relate to multiple purposes. Presenting results with reference to the purposes allows for a clearer link to the discussion, which focuses on the purposes of the study. Presenting with reference to the statistical methods allows results to be shown more completely and with less bias.

See section 3.6.3 for the purpose of each of chosen the statistical methods. See Appendix 13 for the raw data taken directly from participants and Appendix 14 for the same raw data condensed into the form which was used in the statistical analysis.

4.1 Graphical Comparison of Groups

Here the key results are presented in a visual format to give a better overview and understanding of how the three groups compared on Demographics, Performance Times, TLX Demand, TLX Performance and SUS. The three groups are <35 (under 35 years of age), 50+ (over 50 years of age) and DNC (Did Not Complete all four tasks).

Descriptive statistics were used (means and range) to create the graphs seen below, these results do not show statistically significant effects; the following sub sections give the results of deeper statistical analyses.

All Following Graphs use the key seen in Figure 4.1.

4.1.1 Demographics

Key Demographics separating each of the three groups are Age, Gender, Confidence with Technology and Total ICF. There were similar numbers in each groups as seen in Figure 4.1, 15 in <35, 12 in 50+ and 13 participants who started in DNC with 6 participants completing the experiment.

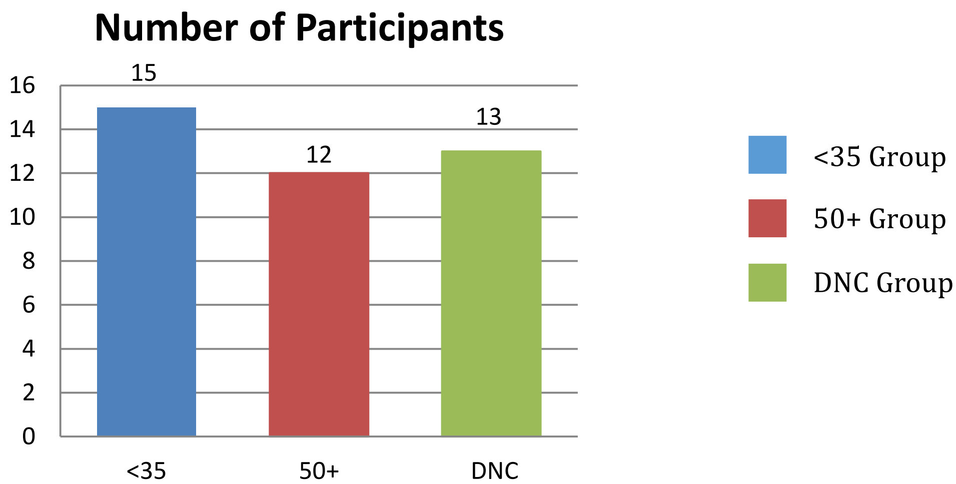


Figure 4.1: Number of Participants in each group (left), key to graph (right).

It can be seen in Figure 4.2 that the Mean Ages are very similar between 50+ and DNC, where <35 is considerably less than both of the other groups. Standard deviations for Age were 4.26

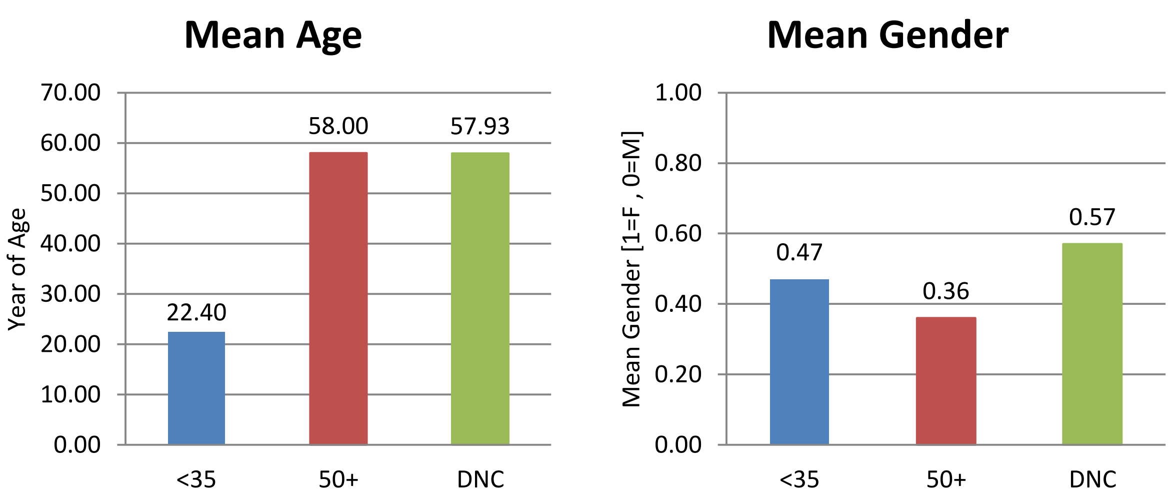


Figure 4.2: Mean age (left) and gender (right) of groups

for <35, 5.66 years for 50+ and 12.55 for DNC.

Gender has a close to even balance for all groups <35 and DNC, with a slight skew for 50+. As can be seen in Figure 4.2, where Male is coded as 0 and Female is coded as 1. DNC is slightly skewed towards female, while <35 is slightly skewed towards Male. 50+ has the greatest imbalance with a skew towards Male, with a mean of 0.36. Standard deviations for Gender were 0.52 for <35, 0.52 for 50+ and 0.51 for DNC.

Mean Confidence with Technology seen in Figure 4.3 was self reported as a score from 0 to 10 where ten was very confident with using new technology and zero was very low confidence. Groups 50+ and DNC both gave mean scores of near 5 while <35 gave a score of 7.40. Standard deviations for Confidence were 1.35 years for <35, 1.68 years for 50+ and 2.43 years for DNC.

Mean Total ICF seen in Figure 4.3 was far greater for DNC at 5.93 compared to 0.13 and 0.82 for <35 and 50+ respectivly. Standard deviations for Total ICF were 0.35 for <35, 1.33 for 50+ and 2.40 for DNC. Group DNC had a lowest ICF level of 3, which three participants had, this means that only participants with an ICF of 2 or less were able to complete all tasks.

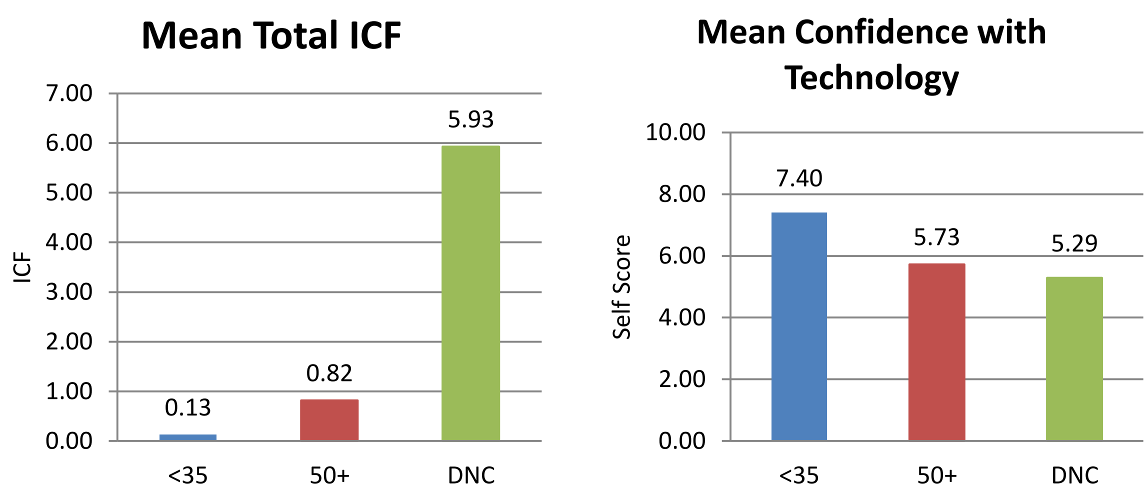


Figure 4.3: Mean ICF (left) and Mean Confidence with using technology (right) of groups

The reported Mean SUS for <35 was the highest at 3.01, followed by DNC at 2.53 and 50+ at 2.21, as seen in Figure 4.4. Where an average score is considered to be 3.4 (Sauro, 2011). Interestingly DNC gave a mid level of usability despite not being able to complete all tasks and (seen in later in this subsection) consistantly having a slower Performance Time. Standard deviations for Total ICF were 0.98 for <35, 0.80 years for 50+ and 0.76 for DNC.

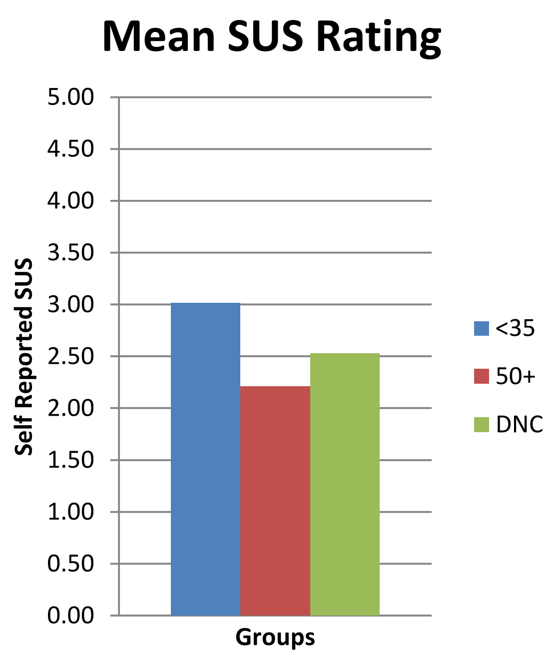


Figure 4.4: Mean System Usability Scale rating of groups

4.1.2 Task Performance Times

Mean Task Performance Time for Task 1 Chair is seen in Figure 4.5. Group <35 had the fastest Mean Performance Time starting at just under 1 minute on Attempt 1 and showing a plateau from Attempt 2 onwards at approximately 8 seconds.

Group 50+ had a slower rate of improvement beginning with a mean Performance Time of approximately 3:30 minutes. Mean Performance Time decreased consistently to Attempt 3 with a mean of 20 seconds by Attempt 4.

Group DNC began at a mean Performance Time of 3:54 minutes, similar to 50+. However DNC's learning curve was not as fast as 50+ having a mean Performance of 1:22 by Attempt 4.

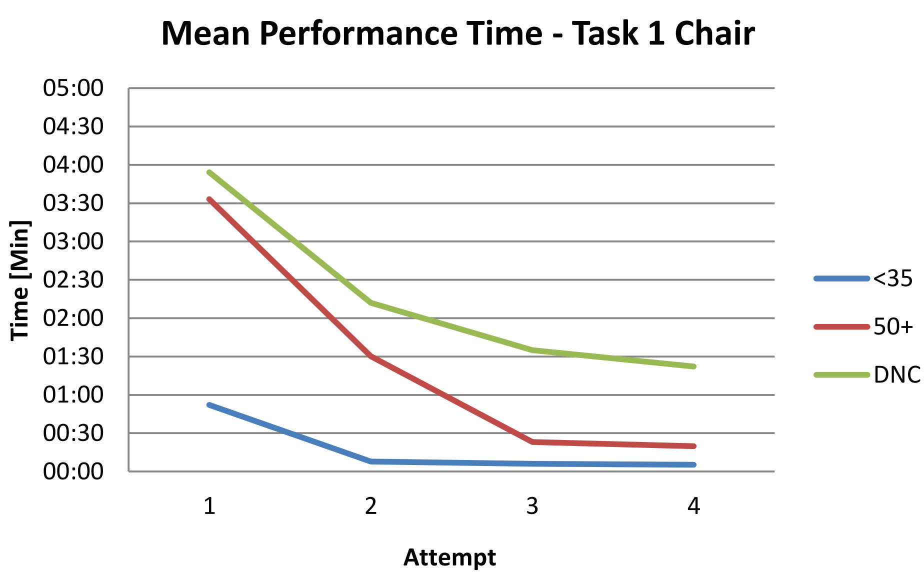


Figure 4.3: Mean performance times of groups for Task 1

Mean Task Performance Time for Task 2 Light is seen in Figure 4.6. This task had the fastest Performance Times of all four tasks.

Groups <35 and 50+ were both comparable with mean Performance times of 16 seconds and 39 seconds respectively for Attempt 1 and 6 seconds and 3 seconds respectively for Attempt 4.

Group DNC showed a similar rate of improvement to the other groups but with slower mean Performance Time. Attempt 1 had a mean time of 2:23 and Attempt 4 a time of 1:05.

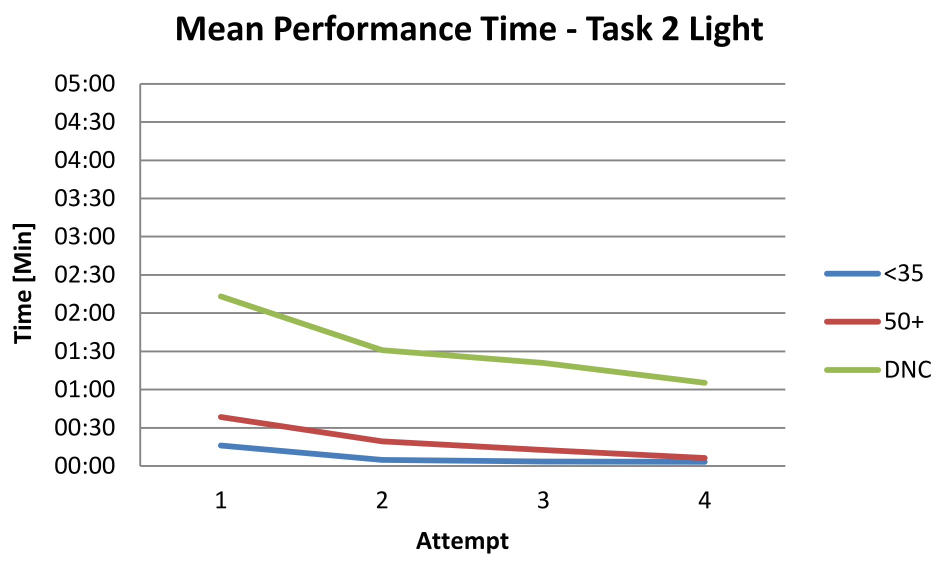


Figure 4.4: Mean performance times of groups for Task 2

Mean Task Performance Time for Task 3 Clock is seen in Figure 4.7. The three groups all have different rates of improvement with <35 being the fastest followed by 50+ and DNC as the slowest.

It is noted that only 4 participants in DNC completed Task 3 Clock, represented by the dashed line. This makes comparisons with the other two groups less reliable.

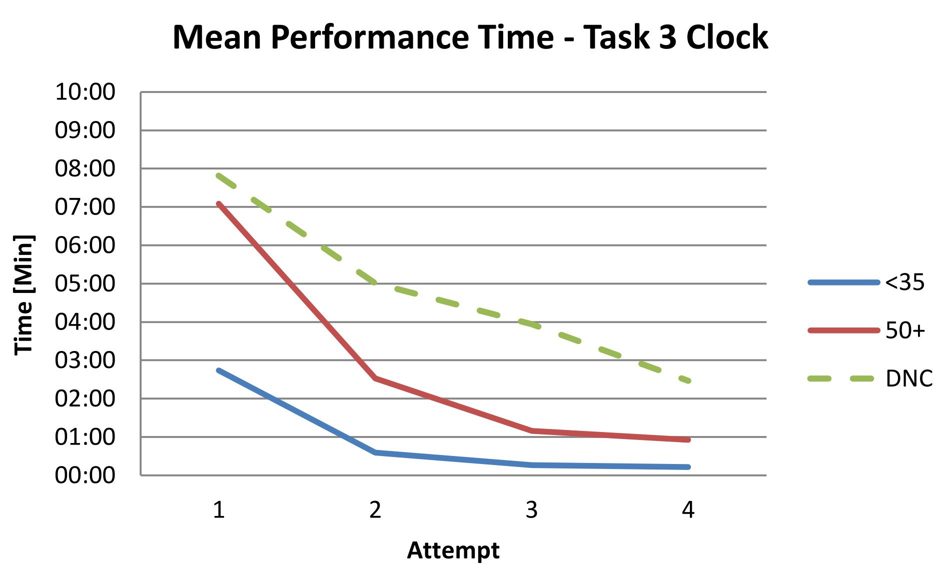


Figure 4.5: Mean performance times of groups for Task 3

Mean Task Performance Time for Task 4 Cog is seen in Figure 4.8. Group DNC only had a single participant and they ended after Attempt 2.

Groups <35 and 50+ had similar rates of learning, with <35 having better Mean Performance Times across all attempts compared to 50+. Group <35 mean Performance Time for Attempt 4 was 0:37 and 1:39 for 50+.

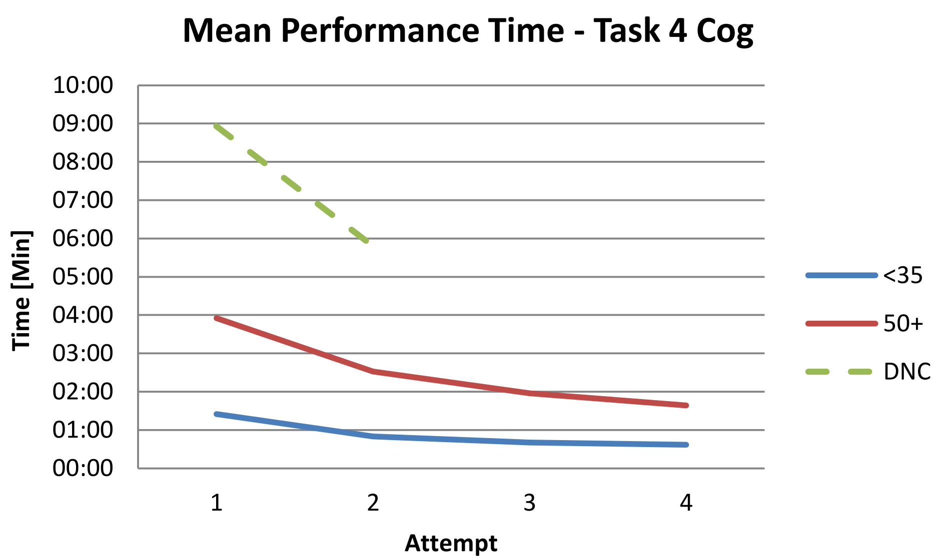


Figure 4.6: Mean performance times of groups for Task 4

The mean Performance Times for Attempt 4 Task 4 Cog can be compared to the summed mean Performance Times for Attempt 4 in the previous three tasks, as seen in Table 4.1. Comparing these times allows a rough indication of the effects of cognitive loading after practice.

The difference between these are 6 seconds for <35 and 18 seconds for 50+. Assuming that non crucial actions such as moving fingers between buttons are negligible between <35 and 50+, the results in Table 4.2 shows that the effects of cognitive loading as applied in this study were consistent between both groups.

The difference between the summed performance times from Task 1, 2 and 3 compared to performance time for Task 4 of <35 compared to 50+ was 3%. Both groups' respective performance times were an order of magnitude different. This suggests that both groups' performances were affected equally by a net increase in time (mean of 17 seconds) despite <35 which had a statistically significantly faster performance in the majority of task attempts.

It is unknown whether this decreased performance for both groups is primarily due to a slowed cognitive recall of the task or delayed movement motor output due to the distraction. From observation of the participants it appeared that they would pause from the task to listen to the distracter word and then continue with the task while replying to the distracter word.

Table 4.1: Performance times for groups on Attempt 4 over all Tasks

Comparison of Performance Times of Attempt 4			
Task	<35 Mean Time Attempt 4	50+ Mean Time Attempt 4	Difference between <35 and 50+
1	0:05	0:20	0:15
2	0:03	0:06	0:03
3	0:13	0:55	0:42
Sum 1, 2, 3	0:21	1:21	1:00 (25.9%)
4	0:37	1:39	1:02 (37.4%)
Difference Sum and T4	0:16 (76.2%)	0:18 (22.2%)	0:02 (3.0%)

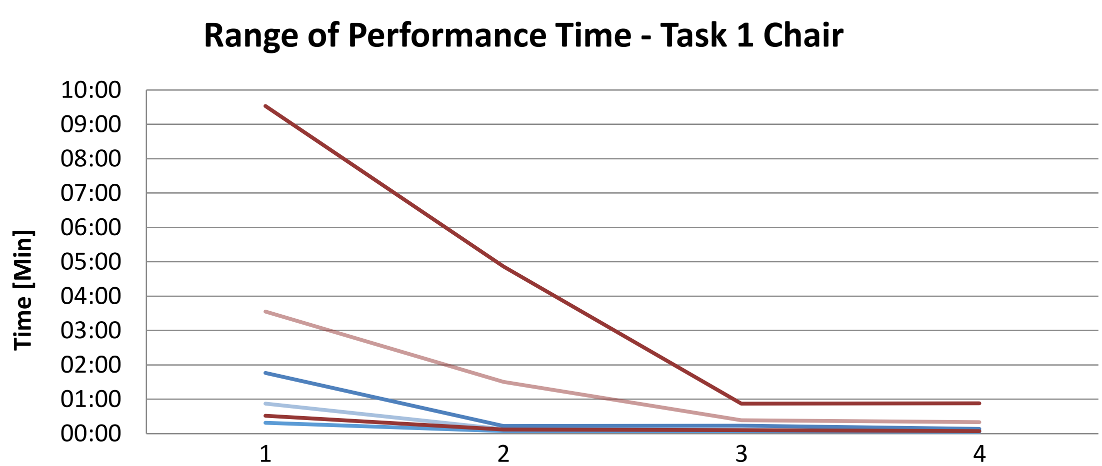
4.1.3 Range of Performance Time

The range of mean Performance Times for each task for groups <35 and 50+ can be seen in Figures 4.9 to 4.12. These graphs are shown to highlight the relative variation in performance between the two groups.

It can be seen in all tasks that 50+ has a greater range than <35, with the greatest discrepancy on Attempt 1 for all tasks.

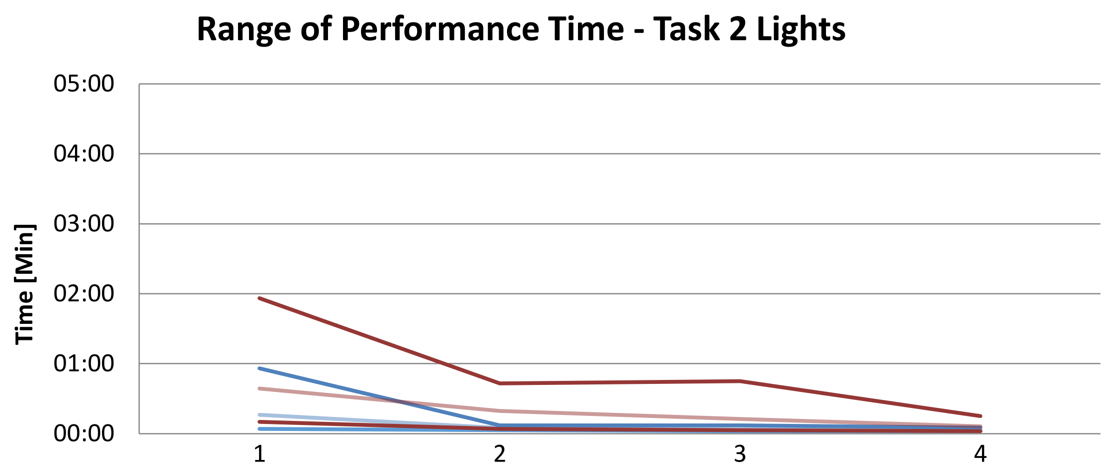
Both Task 3 and 4 had an increase in range for Attempt 4 compared to previous attempts. The max range for 50+ increased compared to the max range for 50+ in Attempt 3. However the mean Performance Time for 50+ does not increase between Attempts 3 and 4.

From the raw data there was only one participant in Task 3 whose Performance Time increased and another single (but different) participant whose time increased in Task 4. It is noted that these participants were not statistically considered to be outliers.



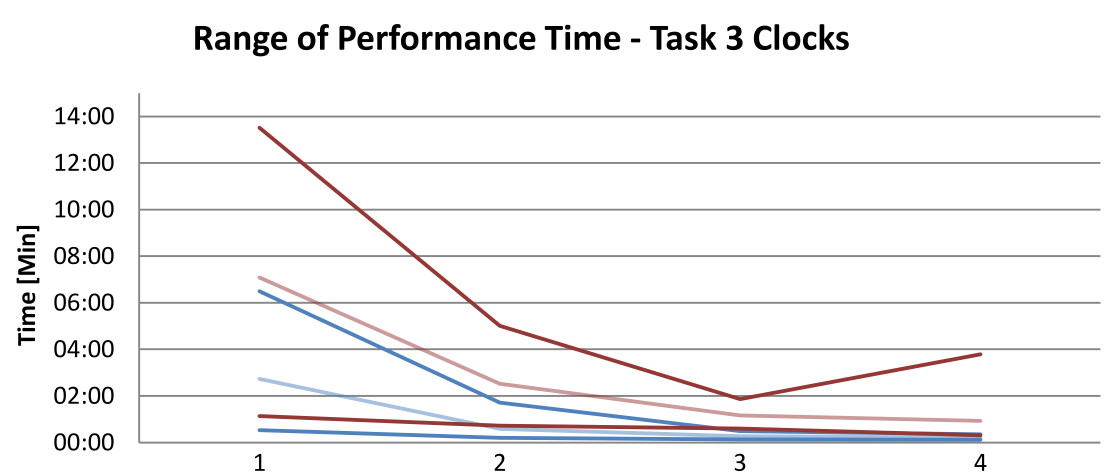
	1	2	3	4
<35 Max	01:46	00:13	00:14	00:08
<35 Min	00:19	00:04	00:04	00:03
<35 Mean	00:52	00:08	00:06	00:05
50+ Max	09:32	04:52	00:52	00:53
50+ Min	00:31	00:07	00:06	00:04
50+ Mean	03:33	01:30	00:23	00:20

Figure 4.7: Performance range in Task 1



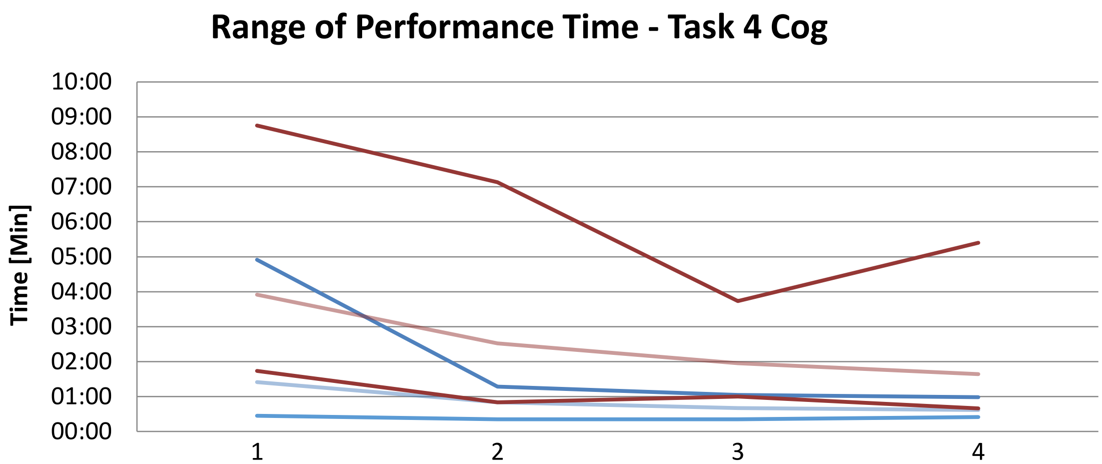
	1	2	3	4
<35 Max	00:56	00:07	00:07	00:05
<35 Min	00:04	00:03	00:02	00:02
<35 Mean	00:16	00:05	00:04	00:03
50+ Max	01:56	00:43	00:45	00:15
50+ Min	00:10	00:04	00:03	00:02
50+ Mean	00:39	00:19	00:13	00:06

Figure 4.8: Performance range of <35 and 50+ in Task 2



	1	2	3	4
<35 Max	06:30	01:43	00:29	00:21
<35 Min	00:32	00:12	00:08	00:07
<35 Mean	02:44	00:35	00:16	00:13
50+ Max	13:31	05:01	01:52	03:47
50+ Min	01:08	00:43	00:36	00:19
50+ Mean	07:05	02:32	01:09	00:55

Figure 4.9: Performance range of <35 and 50+ in Task 3



	1	2	3	4
<35 Min	04:55	01:17	01:03	00:59
<35 Min	00:27	00:21	00:21	00:25
<35 Mean	01:25	00:50	00:40	00:37
50+ Max	08:45	07:08	03:44	05:24
50+ Min	01:44	00:50	01:00	00:40
50+ Mean	03:55	02:31	01:58	01:39

Figure 4.10: Performance range of <35 and 50+ in Task 4

4.1.4 Task TLX Demand and TLX Performance

For all four tasks <35 reported better performance and less demand. There were no distinctive differences between changes (that is the slope of the lines) between the groups on the various tasks, as seen in Figures 4.13 to 4.16.

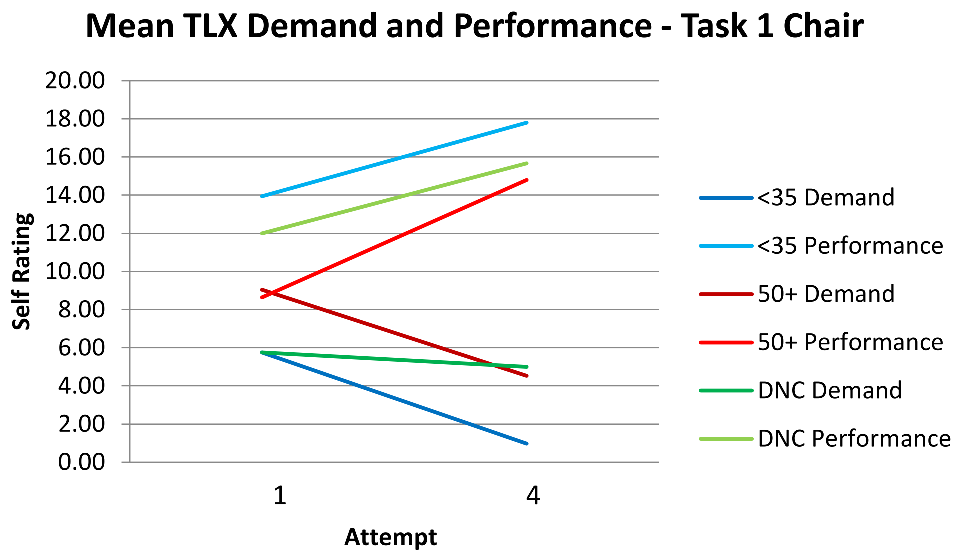


Figure 4.11: Mean demand and performance of groups on Task 1

Task 2 Light was considered to be the least demanding task in general whereas Tasks 3 and 4 were considered the two most demanding.

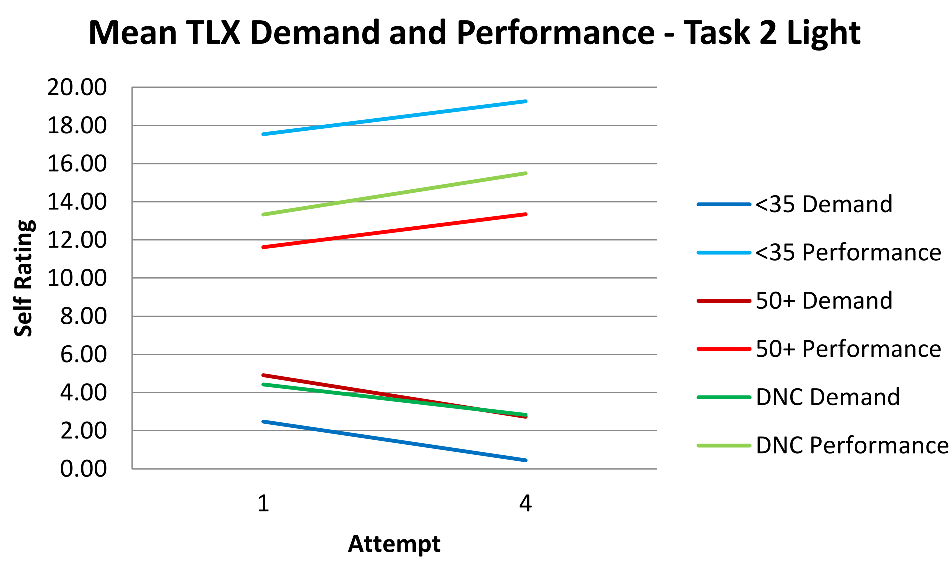


Figure 4.12: Mean demand and performance of groups on Task 2

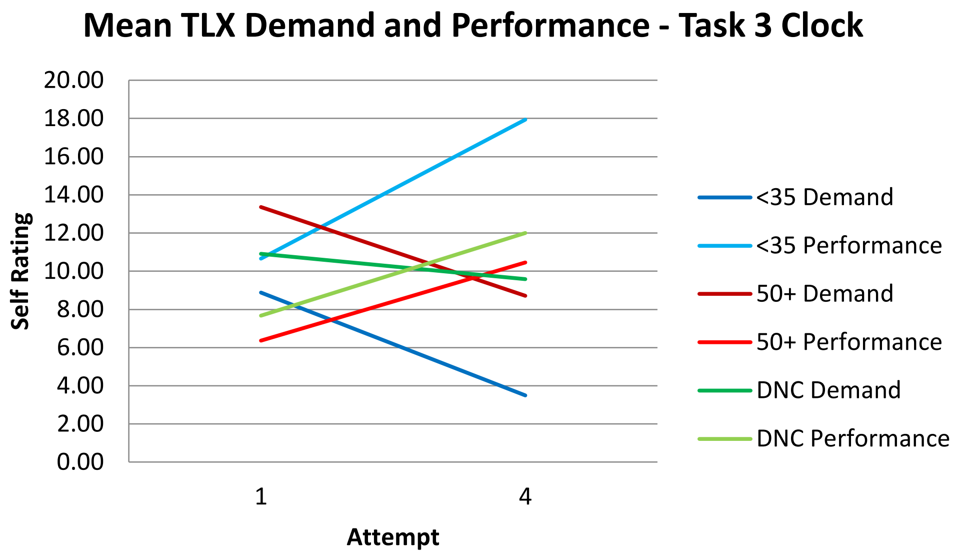


Figure 4.13: Mean demand and performance of groups on Task 2

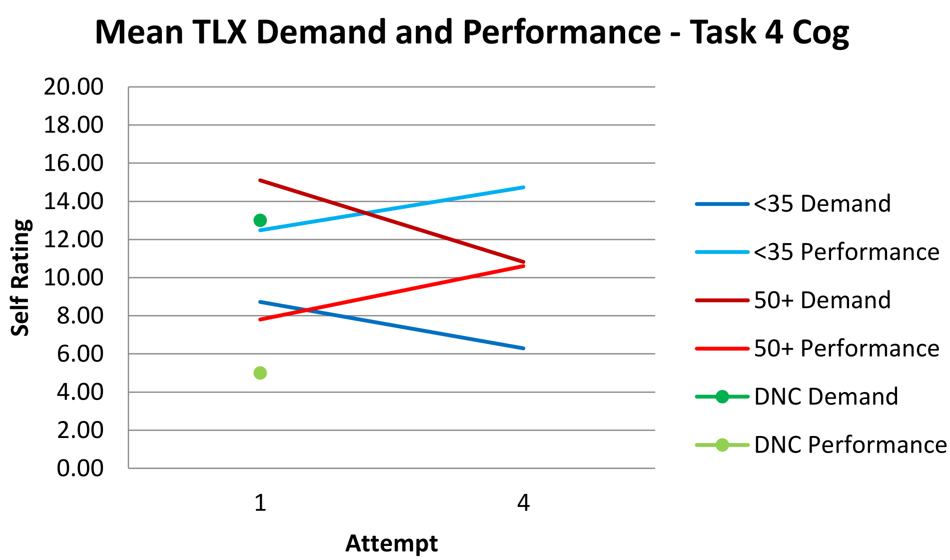


Figure 4.14: Mean demand and performance of groups on Task 2

4.2 Independent T-test - Group Comparisons

Independent T tests were conducted to compare all measured variables between the three groups. Comparisons to DNC group became weaker as the participants in the DNC groups dropped out of the experiment.

The following sub sections give the results for the Independent T tests. Only variables which had a statistically significant difference between the two compared groups are presented.

4.2.1 Group <35 compared to DNC

The variables that had a statistically different mean between groups <35 and DNC are shown in Table 4.2. Presented are descriptive statistics of <35 and DNC for each particular variable and the relevant results from the T-test, significance value (p value) and 95% confidence intervals.

Relevance of comparing <35 and DNC began to fade as the DNC participants begin to drop from the experiment. As can be seen in Table 4.3 the number of participants (column N) decreased as Task 1 Chair began. Only two participants from DNC are present for the majority of Task 3 Clock and only one managed a single attempt at Task 4 Cog.

Groups <35 and DNC had significant differences in all demographic areas apart from gender. They had differences on some aspect of each task but not the majority of attempts or TLX measures in any.

These results show that many variables contribute to the difference in performance and reported usability. DNC had greater impairment and an older mean age, as well as requiring greater assistance and being less confident with technology. This highlights the need the 50+ group to act as a control primarily for the Age variable.

Table 4.2: T-test results for comparison of <35 and DNC

Variable	Group	Descriptive Statistics			T-test for Equality of Means			
		N	Mean	Std. Deviation	Sig. (2-tailed)	Mean difference	95% Confidence Interval of the Difference: Lower	Upper
Demographics								
Assistance given	<35	15	0.13	0.35	0	-0.79	-1.04	-0.54
	DNC	14	0.92	0.28				
Age	<35	15	22.4	4.29	0	-30.83	-39.83	-21.83

	DNC	14	53.23	16.37				
Experience in wheelchairs	<35	15	0	0	0	-0.77	-1	-0.54
	DNC	14	0.77	0.44				
Confidence with new technology	<35	15	7.4	1.35	0.008	2.17	0.63	3.71
	DNC	14	5.23	2.52				
Total ICF	<35	15	0.2	0.41	0	-7.34	-8.7	-5.98
	DNC	14	7.54	2.54				
ICF Specifics (0-4 scale)								
PsycMotor	<35	15	0.07	0.26	0	-2.01	-2.74	-1.28
	DNC	14	2.08	1.19				
Movement	<35	15	0	0	0	-1.77	-2.38	-1.16
	DNC	14	1.77	1.01				
Audio	<35	15	0	0	0.018	-0.38	-0.69	-0.08
	DNC	14	0.38	0.51				
Tactile	<35	15	0	0	0.025	-0.62	-1.14	-0.09
	DNC	14	0.62	0.87				
Skills	<35	15	0.07	0.26	0	-2.24	-2.87	-1.61
	DNC	14	2.31	1.03				
Task attempts (seconds)								
Attempt 1 Task 1 Chair	<35	15	52.13	27.31	0	-181.7	-263.33	-100.07
	DNC	6	233.83	150.61				
Attempt 3 Task 1 Chair	<35	15	6.13	2.7	0	-88.03	-123.16	-52.91
	DNC	6	94.17	67.58				
Attempt 1 Task 3 Clock	<35	15	163.93	109.72	0.004	-257.4	-418.84	-95.96
	DNC	3	421.33	178.11				
Attempt 3 Task 3 Clock	<35	15	16.07	6.75	0	-284.43	-297.91	-270.96
	DNC	2	300.5	20.51				
Attempt 1 Task 4 Cog	<35	15	84.93	73.25	0	-514.07	-676.32	-351.82
	DNC	1	599	NA				
TLX (0-20 scale)								
T2 Lights TLX Start Performance	<35	15	17.53	2.64	0.046	3.13	0.06	6.21
	DNC	5	14.4	3.44				
T3 Clock TLX End Demand	<35	15	3.5	3.3	0.007	-7.63	-12.86	-2.39
	DNC	2	11.13	2.65				
T3 Clock TLX	<35	15	17.93	2.94	0.048	4.93	0.05	9.82

End Performance	DNC	2	13	4.24				
TLX Specifics (0-4 scale)								
TLX Start Physical Demand Task 1	<35	15	0.67	0.72	0.02	-0.83	-1.52	-0.14
	DNC	6	1.5	0.55				
TLX Start Temporal Demand Task 1	<35	15	4.8	4.95	0.108	3.47	-0.83	7.77
	DNC	6	1.33	0.52				
TLX End Physical Demand Task 1	<35	15	0.36	0.61	0.019	-1.64	-2.97	-0.3
	DNC	5	2	2.35				
TLX End Temporal Demand Task 1	<35	15	0.83	1.38	0.067	-1.77	-3.68	0.14
	DNC	5	2.6	2.7				
TLX Start Temporal Physical Task 2	<35	15	0.6	0.74	0.024	-2	-3.71	-0.29
	DNC	5	2.6	3.05				
TLX End Temporal Mental Task 3	<35	15	4.33	4.13	0.003	-10.67	-17.18	-4.15
	DNC	2	15	2.83				
TLX Start Frustration Task 4	<35	15	5.99	5.17	0.029	-13.01	-24.45	-1.56
	DNC	1	19	NA				
SUS (1-5 scale)								
Would need support	<35	15	1.67	1.11	0.001	-2.33	-3.5	-1.17
	DNC	6	4	1.27				

4.2.2 Group 50+ compared to DNC

The variables with a statistically different mean between groups 50+ and DNC are shown in Table 4.3.

There were fewer statistically significant differences between 50+ and DNC than for <35 and DNC.

The significant demographic variables of Experience in wheelchairs, Total ICF and Tasks not completed are expected due to the nature of the two groups. The other significant variables show no clear conclusions as the difference in number of participants between each group is considerable.

From these results it can be concluded that there were no notable differences in task performance times, TLX or SUS measures between 50+ and DNC.

The key differences between 50+ and DNC remain that DNC has significantly higher ICF measurers and they were not able to complete all of the tasks.

There were no notable differences in task performance times, TLX or SUS measures between 50+ and DNC. This suggests that Age and Impairment have similar affects on task performance and usability however it is important to remember that group DNC had a decreasing number of participants over time and that group DNC did not complete the experiment whereas all of 50+ did.

It can be concluded that Age has a similar negative effect on performance and usability comparable to impairment provided the impairment does not prevent the task from being possible. It is likely that if a certain Age provides an equal level of inability as a level of impairment it does not diminish a person’s perseverance to complete the task as the impairment does.

Table 4.3: T-test results for comparison of 50+ and DNC

Variable	Group	Descriptive Statistics			T-test for Equality of Means			
					Sig. (2-tailed)	Mean difference	95% Confidence Interval of the Difference	
		N	Mean	Std. Deviation			Lower	Upper
Demographics								
Experience in wheelchairs	50+	11	0.17	0.39	0.001	-0.60	-0.95	-0.26
	DNC	14	0.77	0.44				
Total ICF	50+	11	1.67	2.15	0.000	-5.87	-7.82	-3.92
	DNC	14	7.54	2.54				
ICF Specifics (0-4 scale)								
PsycMotor	50+	11	0.42	0.67	0.000	-1.66	-2.46	-0.86
	DNC	14	2.08	1.19				
Movement	50+	11	0.08	0.29	0.000	-1.69	-2.31	-1.06
	DNC	14	1.77	1.01				
Skills	50+	11	0.50	0.67	0.000	-1.81	-2.53	-1.09
	DNC	14	2.31	1.03				
Task Attempts (seconds)								
Attempt 3 Task 3 Clock	50+	11	72.57	26.99	0.000	-227.93	-272.04	-183.82
	DNC	3	300.50	20.51				
TLX (0-20 scale)								
TLX Start Temporal Demand Task 1	50+	11	8.00	5.74	0.002	6.67	3.01	10.33
	DNC	7	1.33	0.52				
TLX End Physical Demand Task 2	50+	11	0.38	0.65	0.016	-1.62	-2.88	-0.35
	DNC	6	2.00	1.87				

4.2.3 Group <35 compared to 50+

The variables that had a statistically different mean between groups 50+ and DNC are shown in Table 4.5. The layout of this table is slightly different from those above; here variables are given under their related task, demographics or SUS as opposed to assessments. Descriptive statistics of <35 and 50+ for each particular variable and the relevant results from the T-test are presented.

There are considerably more variables with statistically significant mean differences between <35 and 50+ than with DNC. This is due primarily to the decreasing number of participants in DNC as the experiment progressed.

Performance Times differ between <35 and 50+ in every task on almost every attempt (15 of 16 attempts). The difference in Performance Times decreases between the groups as more attempts are made; the plateau being approached may be statistically different but may not be relevant for practical performance.

Difference in TLX Demand and TLX Performance differ in all tasks but occur most in Task 2 Lights and Task 4 Cog where there are differences in both the End and Start measurers for TLX Demand and TLX Performance. Task 1 Chair shows the least difference with only TLX Start Performance and TLX End Performance. Task 3 Clock shows statistical difference in TLX Start Demand, TLX End Demand and TLX End Performance.

A table showing the breakdown of TLX Demand into TLX Specifics can be seen in Table 4.4. Differences between the groups are predominantly in the End measures for Tasks 1, 2 and 3, while differences for Task 4 Cog are mostly in the Start Measures.

There are consistent differences in End Mental Demand and End Frustration through all four Tasks. Start Temporal Demand is consistent over all four Tasks and Start Temporal Demand is consistent over Tasks 2, 3 and 4.

Table 4.4: Measured variables compared between <35 and 50+ with a significant difference

		Task 1	Task 2	Task 3	Task 4
Start	Mental Demand				✔
	Temporal Demand		✔	✔	✔
	Effort		✔		✔
	Frustration		✔		✔
End	Mental Demand	✔	✔	✔	✔
	Temporal Demand	✔	✔	✔	
	Effort		✔	✔	
	Frustration	✔	✔	✔	✔

Table 4.5: T-test results for comparison of <35 and 50+

Variable	Group	Descriptive Statistics			T-test for Equality of Means			
					Sig. (2-tailed)	Mean difference	95% Confidence Interval of the Difference	
		N	Mean	Std. Deviation			Lower	Upper
Demographics								
Assistance	50+	11	0.73	0.47	0.001	0.59	0.26	0.92
	<35	15	0.13	0.35				
Age	50+	11	58.00	5.66	0.000	35.60	31.58	39.62
	<35	15	22.40	4.29				
Confidence with new technology	50+	11	5.73	1.68	0.010	-1.67	-2.90	-0.45
	<35	15	7.40	1.35				
Task 1 Chair Attempts (seconds) and TLX Results (0-20 scale)								
Attempt 1	50+	11	213.09	175.41	0.002	160.96	66.63	255.28
	<35	15	52.13	27.30				
Attempt 2	50+	11	90.27	89.32	0.001	82.54	35.28	129.80
	<35	15	7.73	2.60				
Attempt 3	50+	11	23.18	16.07	0.000	17.05	8.39	25.71
	<35	15	6.13	2.70				
Attempt 4	50+	11	19.73	13.07	0.000	14.42	7.44	21.40
	<35	15	5.30	1.58				
TLX Start Performance	50+	11	8.64	6.42	0.044	-5.30	-10.45	-0.14
	<35	15	13.93	6.19				
TLX End Demand	50+	11	4.52	3.69	0.002	3.56	1.50	5.61

	<35	15	0.97	1.03				
TLX End Mental Demand	50+	11	4.80	3.54	0.001	3.79	1.79	5.79
	<35	15	1.01	1.13				
TLX End Temporal Demand	50+	11	5.60	5.35	0.003	4.77	1.81	7.73
	<35	15	0.83	1.38				
TLX End Frustration	50+	11	3.76	2.87	0.001	3.02	1.42	4.62
	<35	15	0.74	0.80				
Task 2 Lights Attempts (seconds) and TLX results (0-20 scale)								
Attempt 1	50+	11	38.55	31.01	0.021	22.48	3.70	41.26
	<35	15	16.07	14.62				
Attempt 2	50+	11	19.45	15.40	0.001	14.65	6.48	22.83
	<35	15	4.80	1.08				
Attempt 3	50+	11	12.55	12.99	0.013	8.95	2.03	15.86
	<35	15	3.60	1.24				
TLX Start Demand	50+	11	4.91	2.79	0.014	2.43	0.55	4.31
	<35	15	2.48	1.87				
TLX Start Performance	50+	11	11.62	5.40	0.001	-5.92	-9.22	-2.61
	<35	15	17.53	2.64				
TLX Start Temporal Demand	50+	10	3.98	2.00	0.020	2.31	0.41	4.22
	<35	15	1.67	2.41				
TLX Start Effort	50+	10	4.44	3.64	0.027	2.57	0.31	4.83
	<35	15	1.87	1.81				
TLX End Demand	50+	11	2.72	2.96	0.008	2.27	0.67	3.88
	<35	15	0.45	0.59				
TLX End Performance	50+	11	13.35	7.35	0.005	-5.92	-9.86	-1.98
	<35	15	19.27	1.03				
TLX Start Frustration	50+	10	3.20	1.75	0.008	1.80	0.52	3.08
	<35	15	1.40	1.35				
TLX End Mental Demand	50+	11	2.91	2.88	0.004	2.39	0.82	3.97
	<35	15	0.52	0.63				
TLX End Temporal Demand	50+	11	1.96	2.03	0.012	1.51	0.37	2.66
	<35	15	0.45	0.63				
TLX End Effort	50+	11	3.44	4.23	0.016	2.85	0.58	5.12
	<35	15	0.58	0.62				
TLX End Frustration	50+	11	2.58	3.29	0.021	2.13	0.35	3.92
	<35	15	0.45	0.63				
Task 3 Clock Attempts (seconds) and TLX results(0-20 scale)								
Attempt 1	50+	11	424.91	203.75	0.000	260.98	133.21	388.74
	<35	15	163.93	109.72				
Attempt 2	50+	11	151.73	90.03	0.000	116.53	65.54	167.51
	<35	15	35.20	29.15				

Attempt 3	50+	11	69.35	25.77	0.000	53.28	39.01	67.55
	<35	15	16.07	6.75				
Attempt 4	50+	11	55.41	58.85	0.010	42.61	11.36	73.86
	<35	15	12.80	4.60				
TLX Start Demand	50+	11	13.36	5.69	0.040	4.48	0.21	8.74
	<35	15	8.88	4.82				
TLX Start Temporal Demand	50+	11	12.73	6.33	0.002	7.86	3.08	12.64
	<35	15	4.87	5.46				
TLX End Demand	50+	11	8.71	4.47	0.002	5.21	2.08	8.35
	<35	15	3.50	3.30				
TLX End Performance	50+	11	10.45	5.68	0.000	-7.48	-11.00	-3.96
	<35	15	17.93	2.94				
TLX End Mental Demand	50+	11	9.91	5.30	0.006	5.58	1.76	9.39
	<35	15	4.33	4.13				
TLX End Temporal Demand	50+	11	7.25	5.19	0.000	5.85	3.00	8.71
	<35	15	1.40	1.24				
TLX End Effort	50+	11	9.16	3.99	0.039	4.30	0.25	8.35
	<35	15	4.87	5.53				
TLX End Frustration	50+	11	8.53	5.18	0.016	5.13	1.02	9.23
	<35	15	3.40	4.88				
Task 4 Cog Attempts (seconds) and TLX results (0-20 scale)								
Attempt 1	50+	11	235.09	159.42	0.004	150.16	54.20	246.12
	<35	15	84.93	73.25				
Attempt 2	50+	11	151.36	109.01	0.001	101.23	42.98	159.48
	<35	15	50.13	13.41				
Attempt 3	50+	11	117.53	56.75	0.000	77.05	46.18	107.93
	<35	15	40.48	11.60				
Attempt 4	50+	11	98.55	79.78	0.006	61.67	18.97	104.36
	<35	15	36.88	10.46				
TLX Start Demand	50+	11	15.10	3.05	0.000	6.38	3.61	9.14
	<35	15	8.72	3.58				
TLX Start Performance	50+	11	7.80	3.46	0.016	-4.69	-8.41	-0.96
	<35	15	12.49	5.19				
TLX Start Mental Demand	50+	11	15.69	3.09	0.005	4.21	1.42	6.99
	<35	15	11.49	3.61				
TLX Start Temporal Demand	50+	11	14.73	4.15	0.000	7.53	3.79	11.26
	<35	15	7.20	4.83				
TLX Start Effort	50+	11	16.00	2.72	0.001	5.80	2.53	9.07
	<35	15	10.20	4.69				
TLX Start Frustration	50+	11	13.96	5.18	0.001	7.97	3.73	12.21
	<35	15	5.99	5.17				
TLX End Demand	50+	11	10.83	3.28	0.007	4.53	1.34	7.71

	<35	15	6.30	4.27				
TLX End Performance	50+	11	10.60	2.88	0.008	-4.13	-7.09	-1.18
	<35	15	14.73	4.04				
TLX End Mental Demand	50+	11	12.51	4.21	0.007	4.91	1.45	8.36
	<35	15	7.60	4.22				
TLX End Frustration	50+	11	10.02	4.01	0.001	5.82	2.74	8.90
	<35	15	4.20	3.57				
SUS (1-5 sclae)								
SUS Average	50+	11	2.21	0.80	0.036	-0.80	-1.55	-0.06
	<35	15	3.01	0.98				
SUS System Complexity	50+	11	4.31	0.78	0.006	1.38	0.44	2.31
	<35	15	2.93	1.33				
SUS Support for Use	50+	11	3.64	1.29	0.000	1.97	1.00	2.94
	<35	15	1.67	1.11				
SUS Amount to Learn	50+	11	3.89	1.04	0.025	1.29	0.18	2.41
	<35	15	2.60	1.55				

4.3 Independent T-test – ICF Specifics Affecting Completion

An independent T-test was done similar to those in section 4.2.0, here the ICF Specifics are compared between those participants who completed all the tasks and those who did not (groups <35 and 50+ combined compared to DNC). This was to done to evaluate which of the IC measures had the greatest impact on performance.

From the results of the T-test shown in Table 4.6, it can be seen that there is a significantly different mean between participants who completed and those who did not complete for all ICF Specifics apart from Audio Perception and Tactile Perception.

It is noted that the ICF measures are not wholly independent, therefore weakening the analysis. However useful information can still be gained particularly in conjunction with other analyses done in this study.

Table 4.6: T-test results comparing ICF Specifics of DNC to 50+ and <35

ICF Specific	Group	Descriptive Statistics			T-test for Equality of Means			
					Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
		N	Mean	Std. Deviation			Lower	Upper
PsycMotor Functions	Not Completed	13	2.08	1.188	.000	1.855	1.119	2.591
	Completed	27	0.22	0.506				
Voluntary Movements	Not Completed	13	1.77	1.013	.000	1.732	1.117	2.347
	Completed	27	0.04	0.192				
Audio Perception	Not Completed	13	0.38	0.506	.093	.274	-.050	.597
	Completed	27	0.11	0.320				
Visual perception	Not Completed	13	0.38	0.650	.323	.199	-.215	.614
	Completed	27	0.19	0.396				
Tactile perception	Not Completed	13	0.54	0.660	.019	.501	.098	.905
	Completed	27	0.04	0.192				
Acquiring skills	Not Completed	13	2.31	1.032	.000	2.048	1.402	2.695
	Completed	27	0.26	0.526				

4.4 Stepwise Regression

A stepwise regression analysis was conducted to identify which of the measured variables most contributed to the SUS.

Four models were created from the analysis with a total of four independent variables having a statistically significant effect on SUS used in the final model. Model summaries can be seen in Table 4.7.

TLX Start Demand Task 3 Clock in model 1 was the strongest predictor with an adjusted R squared value of 0.491 that is a 49.1% unique contribution to the variance of SUS.

Other variables were less influential on SUS. The unique contribution of each added variable on the models is seen in the change in adjusted R squared column in Table 4.8. Attempt 1 Task 4 Cog, TLX Start Performance Task 2 Lights and Attempt 2 Task 2 Light add 6.6%, 5.8% and 10.7% contributions to each additional model respectively compare to the 49.1% from TLX Start Demand Task 3 Clock.

Table 4.7: Statistical models predicting SUS

Model Summary – Prediction of SUS				
Model	R Squared	Adjusted R Square	Change in Adjusted R Squared	Independent Variables
1	.511	.491		TLX Start Demand Task 3
2	.592	.557	.066	TLX Start Demand Task 3, Attempt 1 Task 4
3	.661	.615	.058	TLX Start Demand Task 3, Attempt 1 Task 4, TLX Start Performance Task 2
4	.766	.722	.107	TLX Start Demand Task 3, Attempt 1 Task 4, TLX Start Performance Task 2, Attempt 1 Task 2

4.5 Multiple Regression

Three dependent variables were examined using multiple regression analyses; these were TLX End Demand, TLX End Performance and SUS. Models were built using relevant independent variables added in chronological order to predict the dependent variables.

TLX End Demand and TLX End Performance were examined for each of the four tasks separately using the corresponding task Performance Times and TLX Start Demand and Performance measures. SUS was examined independently using TLX Demand, TLX Performance and Attempt 1 Performance Times and Attempt 4 Performance Times from all four tasks

Table 4.8: Comparison of subjective measures through multiple regression

Multiple Regression Statistical Models						
Dependent Variable	R squared	Independent Variables (Change in R squared as variable is added)				
TLX End Demand Task 1	0.742	TLX Start Demand (39.4%)	Attempt 1 (17.2%)	Attempt 2 (17.2%)	Attempt 3 (0.2%)	Attempt 4 (0.2%)
TLX End Demand Task 2	0.831	TLX Start Performance (16.5%)	Attempt 1 (1.5%)	Attempt 2 (5.5%)	Attempt 3 (9.6%)	Attempt 4 (0%)
TLX End Performance Task 2	0.655	TLX Start Performance (66.5%)	Attempt 1 (1.5%)	Attempt 2 (5.5%)	Attempt 3 (9.6%)	Attempt 4 (0%)
TLX End Demand Task 3	0.833	TLX Start Demand (62.8%)	Attempt 1 (0.1%)	Attempt 2 (10%)	Attempt 3 (10%)	Attempt 4 (0.4%)
TLX End Performance Task 3	0.566	TLX Start Demand (19.7%)	Attempt 1 (19.7%)	Attempt 2 (7.2%)	Attempt 3 (9.9%)	Attempt 4 (0.1%)
TLX End Demand Task 4	0.67	TLX Start Demand (64.8%)	Attempt 1 (0.2%)	Attempt 2 (0.9%)	Attempt 3 (0.4%)	Attempt 4 (0.4%)
TLX End Performance Task 4	0.688	TLX Start Performance (59%)	Attempt 1 (0.1%)	Attempt 2 (7.1%)	Attempt 3 (0.2%)	Attempt 4 (0.4%)
SUS	0.565	TLX Start Demand T1 (24%)	TLX Start Demand T2 (0.5%)	TLX Start Demand T3 (31.1%)	TLX Start Demand T4 (0.9%)	
SUS	0.425*	Attempt 1 T1 (12.6%)	Attempt 1 T2 (1.6%)	Attempt 1 T3 (20.5%)	Attempt 1 T4 (7.8%)	

*not over 0.5 but included as referred to in the discussion

Table 4.8 shows the models associated with each of the three dependent variables which have an R squared value of 0.5 or greater. These models do not meet all assumptions of the regression analysis but the reported R squared values are still relevant despite this. The models show the independent variables with the associated change in R squared as a percentage contribution to variance of the dependent variable as each variable was added on, with contributions of 20% or shown in bold.

4.6 Linear Mixed Effects Model

A linear mixed effects model was used to compare fixed effects of Performance Times between the three groups, the four tasks and the four attempts independently of each other.

All fixed effects were significant contributors to the dependent variable Performance Times and had significant differences in terms of time measures. Table 4.9 shows the relevant results from the analysis.

Comparing the groups, <35 completed all attempts on average 122 seconds faster than DNC and 79 seconds faster than 50+. Group 50+ was 43 seconds faster than DNC in completing attempts and was 64% slower than <35.

Task 3 was the longest taking on average 149 seconds to complete across all attempts and groups. Task 4 Cog took 119 seconds, Task 2 Lights took 77 seconds and Task 1 Chair was the fastest to complete in 42 seconds.

For the attempts; Attempt 4 was the fastest with Attempt 3, 2 and 1 taking 9, 39 and 128 seconds more respectively to complete on average across all Task and Groups.

Table 4.9: Results of linear effects model

Estimates of Fixed Effects					
	Parameter	Difference in Time [sec]	Sig.	95% Confidence Interval	
				Lower Bound	Upper Bound
Groups	<35	-122.69	.000	-162.22	-83.16
	50+	-43.89	.036	-84.62	-3.16
	DNC	Base reference			
Tasks	T1 Chair	-41.80	.000	-64.20	-19.39
	T2 Lights	-77.48	.000	-99.94	-55.01
	T3 Clock	30.26	.009	7.58	52.94
	T4 Cog	Base reference			
Attempts	Attempt 1	128.93	.000	106.85	151.01
	Attempt 2	39.09	.001	17.01	61.17
	Attempt 3	9.02	.422	-13.06	31.10
	Attempt 4	Base reference			

4.7 Cox Regression Survival Analysis

This experiment consisted of a total of four tasks each with four attempts making a total of 16 attempts. Fourteen of the forty participants did not complete all 16 attempts, the DNC group. A survival analysis was used to identify the independent variables that correlated with the event of not completing a task. The independent variables used were Age, Total ICF and ICF specifics. Age had no statistically significant effect while Total ICF did. Of the ICF Specifics only Acquiring Skills had a significant effect on the number of tasks completed.

4.7.1 Graphed Survival Function

Figure 1.17 shows the survival time graphed for all participants over the course of the experiment. Survival time is measured in attempts completed. A total of 40 participants began the experiment with a drop to 33 on the first attempt as 7 participants were unable to complete Attempt 1 of Task 1 Chair. Drop offs continue steadily until a remaining 26 participants complete the final 16th attempt, that is Attempt 4 of Task 4 Cog.

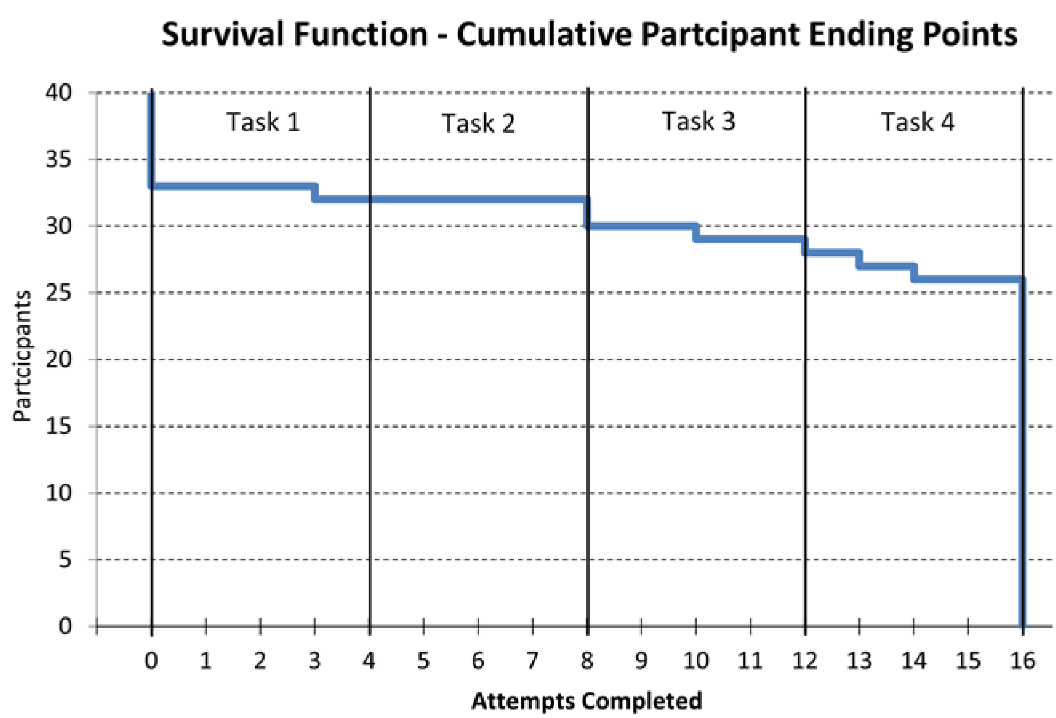


Figure 4.1: Survival Analysis of all participants

From Figure 1.17 it can be seen that the largest drop off of seven participants occurs on the first attempt, if participants are able to complete Attempt 1 Task 1 Chair then they are likely to complete the first two tasks. From the beginning of Task 3 Clock to Attempt 2 Task 4 Cog six participants quit with the remaining 26 participants completing all tasks.

4.7.2 Cox Regression – Age and ICF

Age and ICF were considered the two most important variables to examine in predicting failure to complete. Age, Total ICF and the interaction of the two (Age*Total ICF) were tested. It can be seen in Table 4.10 that neither Age nor Age*Total ICF had a statistically significant effect, while Total did.

Total ICF had a Hazard Ratio[5] of 1.6, meaning that for each point of Total ICF a participant scores they are 1.6 times more likely not to complete all the experiment.

Age and Age*Total ICF are not significant and their Hazard Ratios are approximately 1. This means that there is approximately no rise in likelihood to not complete for an increase in Age or for the interaction of increasing Age and Total ICF.

As this study was unable to find participants who would be considered young and having an impairment (under the age of 35 and with an ICF level of approximately 4 or greater) the effect of Age may be a stronger contributor than indicated here.

Table 4.10: Cox regression Age and ICF

Age and ICF predicting DNC		
	Sig. (p)	Hazard ratio
TotalICF	.021	1.567
Age	.765	1.011
Age*TotalICF	.829	0.999

[5] The hazard ratio is an expression of the chance of hazard event occurring in the treatment as a ratio to the hazard events occurring in the control

4.7.3 Cox Regression – Total ICF and ICF Specifics

Total ICF and the ICF Specifics were examined more closely. A Cox regression was conducted with Total ICF restricted to thresholds at scores of 1 or more, 2 or more, 3 or more and 4 or more with each being used in a separate regression test. From Table 4.11 it is seen that ICF >= 1 had an associated hazard ratio of 73.39, this indicates that participants with a Total ICF level of 1 or greater were approximately 73 times more likely to not complete all the tasks. Similarly participants with an ICF >=2 were 117 times more likely to not complete compared to an ICF < 2, for ICF >= 3 they were 181 times and ICF >= 4 were 825 times more likely to not complete. Where ICF >= 1 did not have a statistically significant effect and the other thresholds did.

Table 4.21: Cox regression ICF

ICF thresholds predicting DNC		
	Sig. (p)	Hazard ratio
ICF >= 1	0.061	73.386
ICF >= 2	0.046	116.961
ICF >= 3	0.042	181.340
ICF >= 4	0.048	525.344

The results of the Cox regression for all ICF Specifics can be seen in Table 4.12. Acquiring Skills was the only measure that had a statistically significant effect, with a hazard ratio of 3.460.

PsycMotor, Voluntary Movements, Visual Perception and Tactile Perception all indicate that an increase of their respective ICF levels will increase the likelihood to not complete all tasks; however none of these measures had a statistically significant effect.

Audio Perception was not statistically significant with a Hazard Ratio of 0.141 (less then 1.0). This hazard ratio was unexpected and suggests that for each point of increase in an ICF level for Audio Perception they are more likely to complete the tasks. The reason may be that there are no levels for Audio Perception greater than 1. Only eight of the 40 participants had a level of 1 for Audio Perception, three of these eight did complete all tasks and they had Total ICF levels of 2, 3 and 7. Thus Audio Perception in comparison to the other ICF Specifics appeared to aid completion as the ICF levels for these three participants were high compared to the other participants who completed all the tasks. This is incorrect as being deaf would not make someone a more adapt user of technology and shows that greater numbers of participants are needed to better identify the effect of these specific ICF measures.

Similar to Audio Perception the effects of Tactile Perception are likely skewed. Only seven of the 40 participants had a level over a zero on this measure, of these six had a level of 1 and the others had a level of 3, likely increasing the affect.

It is clear from the Cox Regression analysis that Acquiring New Skills had the greatest effect on performance but there are non-conclusive results to distinguish how the other specific ICF measures vary in the impact they have on using a device.

Table 4.32: Cox regression ICF predicting DNC

ICF thresholds predicting DNC		
	Sig. (p)	Hazard ratio
PsycMotor Functions	0.178	1.729
Voluntary Movements	0.362	1.512
Audio Perception	0.079	0.141
Visual Perception	0.442	1.917
Tactile Perception	0.117	2.743
Acquiring Skills	0.002	3.460

4.8 Content Analysis of Interview

After completing the tasks and written assessments participants were given an interview and were asked to give general comments on the controller and its use, physical layout, icon clarity, if they made any mistakes (conscious confusion over a decision), lapses (trouble recalling previous actions) or slips (accidently pressing a button they weren't intending to).

The written recordings from these interviews were processed using a basic content analysis method. For each question reoccurring themes in the answers were identified and then tallied for all participants and the three groups. The results from the DNC groups generally only apply to the first two tasks.

The content analysis data based on the recorded notes is shown in Appendix 14.

4.8.1 General Comments

From the general comments the two most reoccurring themes were 'confusion' and 'not intuitive'. The <35 group followed this pattern but with a more even spread they also had a higher number of recommendations. The majority of comments from the 50+ group consisted of 'confusion', 'hard to navigate', 'not intuitive' and 'hard to learn'. Most of the comments of the DNC group were themed around 'confusion' (46% of the total given) with a roughly even spread over the other themes. Results are seen graphically in Figure 4.18 and Table 4.13.

Recommendations given by participants are listed as follow:

- Create more obvious menu themes
- Give the clock a simple on/off toggle
- Make icons clear with text and a picture
- Add a cursor or position indicator
- Reduce the amount of 'scrolling' needed to navigate
- Add an indicator or prompts about what an action would do

The need for a 'back button' was a recommendation made so often it was given a theme. Participants appeared to have trouble navigating to a familiar screen when they became confused.

Praised features included being able to use the joystick to navigate left and right (praised twice) and the looping menus.

Table 4.43: General results of Interview

Content Analysis of General Comments				
Themes	All participants	Group <35	Group 50+	Group DNC
Experienced a fast rate of learning	3	3	0	0
Recommendation given	7	5	1	1
Confusion	26	8	6	12
Inconsistency	2	1	1	0
Dislike	5	2	2	1
Frustration	8	4	0	4
No feedback from controller	5	3	0	2
Back button requested	3	2	0	1
Found hard to learn	6	1	3	2
Trouble navigating	7	0	4	3
Not intuitive to use	13	5	4	4
Visual clarity issue	4	2	1	1
Praised a feature	3	1	1	1

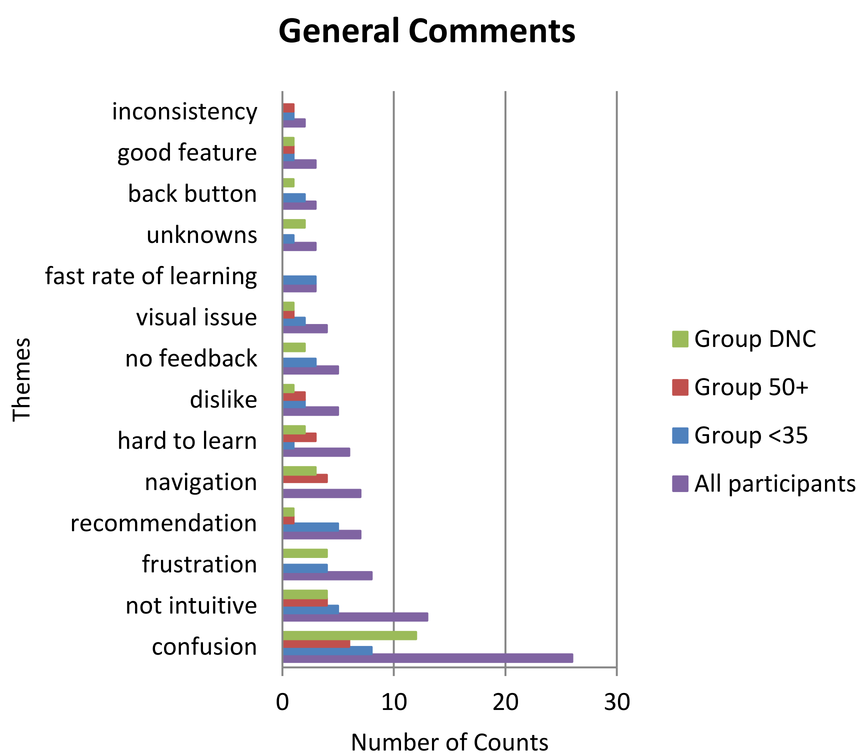


Figure 4.15: Content analysis of general comments from Interview

4.8.2 Physical Layout

All groups generally considered the physical layout to be 'fine'. Complaints ranged across all groups, specific themes were, an uncomfortable length from the joystick to the buttons, that there were too many buttons, and confusion with how the buttons and joystick related to control of the GUI. Results are seen graphically in Figure 4.19 and Table 4.14.

Table 4.54: Physical layouts results of Interview

Content Analysis of Physical Layout Comments				
Themes	**All participants**	**Group <35**	**Group 50+**	**Group DNC**
long distance from joystick to buttons	2	2	0	0
poor link between controls and GUI	2	2	0	0
fine	14	4	8	2
dislike	1	1	0	0
too many buttons	2	2	0	0
Generally poor ergonomics	3	0	3	0
small screen	1	0	0	1
buttons hard to press	2	0	0	2

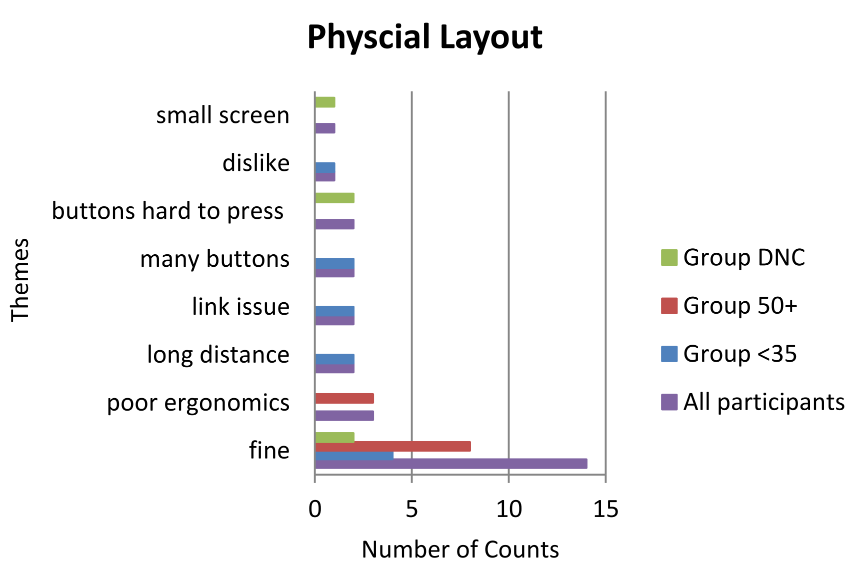


Figure 4.16: Content analysis of physical layout from Interview

4.8.3 Icon Clarity

The majority of feedback from all participants regarding icon clarity was negative. Participants stated that the icons were not intuitive, disliked or that specific tasks icons were not clear. Comparing groups, the majority of <35 found that the icons were not clear or intuitive whereas 50+ were nearly evenly split between 'fine' (acceptable) and a negative response. Results are seen graphically in Figure 4.20 and Table 4.15.

Table 4.65: Icon clarity results of Interview

Content Analysis of Icon Clarity Comments				
Themes	**All participants**	**Group <35**	**Group 50+**	**Group DNC**
Not all intuitive	9	6	1	2
Fine (acceptable)	8	2	6	0
Dislike	2	0	2	0
Chair icons not clear	1	0	1	0
Light icons not clear	3	2	0	1
Clock icons not clear	3	3	0	0
Reflection on screen	1	0	1	0

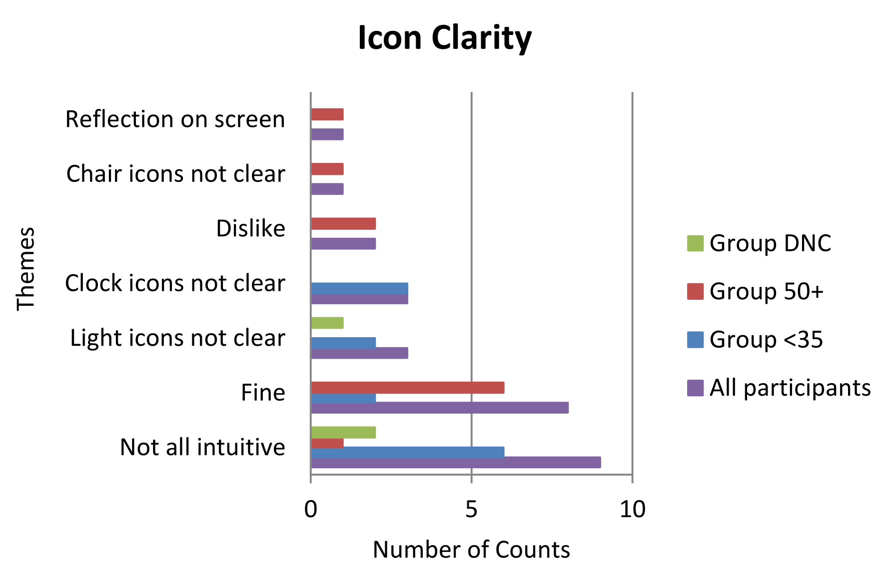


Figure 4.17: Content analysis of icon clarity from Interview

4.8.4 Error Analysis

Reported Mistakes

Mistakes were reported by 15 of 26 participants (57.6%), with more reported in the 50+ group than <35.

The 50+ group were mostly unable to identify in which tasks the mistakes occurred, whereas half of the few mistakes reported for the <35 group occurred on the first attempt of tasks. This resulted in approximately 75% of 50+ and 43% of <35 with confirmed or possible mistakes occurring on one or more of the tasks. Results are seen graphically in Figure 4.21 and Table 4.16.

Table 4.76: Reported mistakes results of Interview

Content Analysis of Reported Mistake Comments				
Themes	**All participants**	**Group <35**	**Group 50+**	**Group DNC**
Yes	9	1	8	
On 1st attempts	3	3	0	
On chair task	1	1	0	
On light task	1	1	0	NA
On cog task	1	0	1	
No	0	0	0	
On clock task	0	0	0	

Figure 4.18: Content analysis of reported mistakes from Interview

Reported Lapses

Lapses were more commonly reported than mistakes with 20 reported for 26 participants (76.9%). Both groups were generally unable to report when lapses occurred. 50+ reported more definitely that there were occurrences of lapses whereas <35 had more definite No's and a small number of Maybe's. This resulted in approximately 64% of 50+ and 58% of <35 with confirmed or possible lapses occurring on one or more of the tasks. Results are seen graphically in Figure 4.22 and Table 4.17.

Table 4.87: Reported lapses results of Interview

Content Analysis of Reported Lapse Comments				
Themes	**All participants**	**Group <35**	**Group 50+**	**Group DNC**
Yes	12	4	8	Na
No	4	9	0	
Maybe	2	2	0	
On cog task	2	1	1	
On 1st attempts	0	0	0	
On chair task	0	0	0	
On light task	0	0	0	
On clock task	0	0	0	

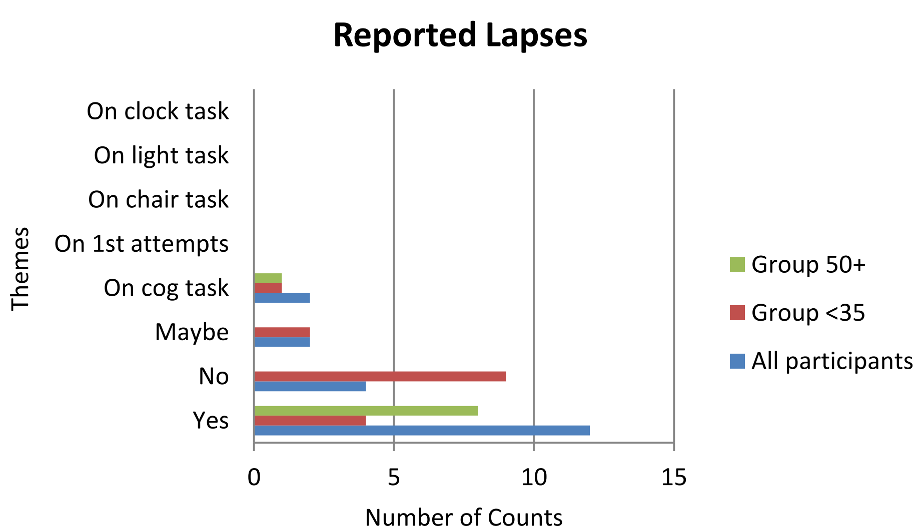


Figure 4.19: Content analysis of reported lapses from Interview

Slips

Most participants reported slips as not occurring. Only 2 of the 26 participants in the 50+ groups reported definite occurrences of slips. Results are seen graphically in Figure 4.23 and Table 4.18.

Table 4.98: Reported slips results of Interview

Content Analysis of Reported Slips Comments				
Themes	**All participants**	**Group <35**	**Group 50+**	**Group DNC**
No	13	6	7	
Yes	2	0	2	NA
Maybe	2	0	2	

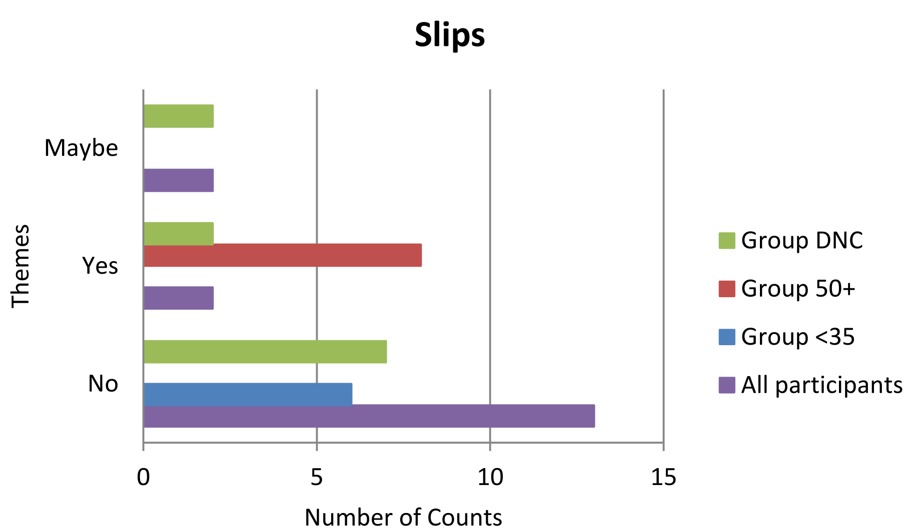


Figure 4.20: Content analysis of reported lapses from Interview

4.10 Implication of Results in Regard to Purposes

4.10.1 Purpose - Identify the effects of impairments on a user's measured performance, and self-reported performance and usability experience

To clarify terms used:

- Impairments were measured using the ICF assessment.
- Performance relates to the time taken to complete an attempt at a task (Performance Time).
- Self-reported performance relates to the TLX assessment, which measures the participant's subjective task demand.
- Performance and usability relates to the SUS assessment and rate of change in Performance Time and TLX.

The statistical assessments evaluating the effects of ICF presented here are the Independent T-tests comparing groups and ICF specifics, Linear Mixed Effects Model, Cox Regression and Content Analysis.

Comparisons between <35 and 50+ groups to DNC are used as the DNC group has a significantly greater Total ICF than the other three groups. The mean Total ICF for DNC is 5.93 compared to 0.13 for <35 and 0.82 for 50+ (shown in Figure 4.3). These differences are statistically significantly different according to the Independent T-test conducted between <35 and 50+ (see Table 4.6).

Affect on Task Completion

From the Cox Regression and T-test that examine ICF Specifics it appears that Acquiring Skills has the greatest affect on a participant's ability to use the controller effectively and therefore performance.

From the Cox Regression survival analysis (Section 4.7) it was determined that a participant with an ICF of 1 or more was 73 times more likely to fail to complete the experiment at some point compared with a participant with an ICF of zero.

The mean total ICF of DNC was near 6 and lowest ICF level in the DNC groups was 3, which three participants had. From the Cox Regression a participant with an ICF of 3 or more was 181 times more likely to not complete the experiment compared to a participant with an ICF of 2 or less.

Looking at the influence of the ICF specifics on group DNC, Acquiring Skills was the only one of the six measures from the Cox Regression analysis, which had a significant effect. Each ICF point indicating a greater impairment for acquiring or learning a new skill gives a 3.5 times greater likelihood that the participant will not compete the experiment. The other ICF measurers should still be considered contributors but were not as independent of other variables as Acquiring Skills in this study.

The Independent T-test comparing means between the DNC participants showed PsycMotor Functions, Voluntary Movements, Tactile perception and Acquiring skills as having statistically significantly different means. PsycMotor Functions, Voluntary Movements and Acquiring skills having the greatest significance.

Effect on Performance

The clearest effect of ICF is that as Total ICF increases the likelihood of a participant not completing the experiment substantially increases and performance reduces. This is shown most clearly in sections 4.2.1 and 4.2.2 where group DNC is compared to the other two groups. Only participants with an ICF of 2 or less were able to complete all tasks.

The Linear Mixed Effects Model and the group comparison T-test show that on the whole participants in the DNC group have a slower performance time compared to the other participants. However for each of the separate attempts few are considered to be statistically significantly slower. It is likely that this poor statistical significance is likely due to the participants in the DNC group dropping from the experiment reducing the N value of the group and therefore reducing statistical strength.

As would be expected participants with higher levels of ICF have worse performance in addition to being less likely to complete the experiment. From the Linear Mixed Effects Model (Table 4.10) DNC has the slowest performance time across all completed attempts by approximately 44 seconds compared to 50+ and 2 minutes compared to <35. This trend of slower Performance Time can be seen graphically in Figures 4.5 to 4.8.

Looking at the task attempts in greater detail the Independent T-test (Table 4.3) shows that there is a statistically significant difference between the Performance Times for <35 and DNC on Attempt 1 and 3 for Task 1 Chair and Task 3 Clock , and Attempt 1 Task 4 Cog. Comparing DNC to 50+ the only significant difference for performance time was on Attempt 3 Task 3 Clock. However there were only 5 participants in group DNC after Task 2.

Effect on self reported Demand, Performance and Usability

Participants with higher ICF generally did not differ significantly on self reported, performance or usability (as seen in Figures 4.13 to 4.16).

From the Independent T-test (seen in Table 4.3) groups <35 and DNC were statistically significantly different on almost all measures. Only three of a possible sixteen measures for TLX Demand and TLX performance were found not to be statistically significant.

Examining the TLX Specifics the majority of significant differences were for Temporal Demand with these occurring for start and end of Task 1 Chair and end for Task 2 Lights and Task 3 Clock. Comparing 50+ and DNC (Table 4.4.) the only significantly different means were the start temporal demand for Task 1 and end physical demand for Task 2.

SUS had very few significant differences between DNC and either of the other groups, only the SUS Specific 'would need support' had a significant difference between <35 and 50+.

All three groups gave an SUS mean near 2.5 with DNC giving a mean of 2.53 despite having worse Performance Times and all 14 participants not completing the experiment.

Six of the 14 participants in DNC undertook the close out interview and contributed to the content analysis. These participants reported a high level of confusion and general dislike of the controller.

4.10.2 Purpose – Determine the effects of cognitive loading on a user's measured and reported performance

Cognitive loading, which was applied in this experiment as a word association task, appeared to have an effect on the participants Performance Time of participants and a greater effect on reported performance. However this effect was similar or not as strong as having to learn to use the controller such as in Task 1 and clearly weaker than struggling with the difficult Task 3.

Task 4 was the final task performed by participants; it consisted of the objectives of the previous three tasks being repeated while performing a word association task. No participants from DNC completed Task 4 so the results are only relevant for <35 and 50+.

Effect on Performance

From the Linear Mixed Effect Model (seen in Table 4.10) it was determined that Task 4 Cog in general had the second slowest Performance time for all groups with Task 3 Clock slower to complete and Task 1 and 2 faster to complete.

Examining the mean time taken to complete the various tasks on the attempts (shown in Table 4.10) group <35 took 16 seconds for Attempt 4 compared to 21 seconds for the sum of the previous three tasks, a difference of 4 seconds. Group 50+ took 1:39 minutes to complete Task 4 Cog compared to a sum of 1:21 minutes by <35, a difference of 18 seconds. Assuming that a near fastest performance time has been reached by Attempt 4 on all tasks then the difference in time between Task 4 Cog and the sum times for Attempt 4 on the other three tasks may roughly represents the slowing effect of the cognitive loading.

Effect on self reported Demand, Performance and Usability

Means were compared between groups on Task 4 Cog for Performance Times, TLX Demand and TLX Performance. There were no significant differences between means for 50+ and DNC. Groups <35 and DNC had two significant measures for Task 4; these were Attempt 1 Performance Time and the Start Frustration (refer to glossary for definition). Groups <35 and 50+ had a significant difference for all four Attempts, Start and End Performance, Start and End Demand, and most Specific TLX Start and End Demands. Group <35 and 50+ had significant differences on measures in all four tasks that were similar to those in Task 4 Cog, Task 4 was unique in that it did have the strongest differences between <35 and 50+ in TLX Demand compared to the other tasks.

Task 4 shows a smaller difference in change between TLX Start and End Demand, and TLX Start and End Performance compared to Task 1 and 3 as seen in Figures 4.13 to 4.16. It is larger that Task 2 Lights, which appears to be the easiest Task based on Performance Times and self-reporting.

From the Stepwise regression analysis Task 4 Cog has a small contribution to SUS, Attempt 1 Task 4 contributes 6.6% to the total variance of SUS compared to 51.1% for the strongest contributor TLX Start Demand Task 3 Clock.

Multiple regressions showed that TLX End Performance and TLX End Demand for Task 4 Cog were most influenced most by their respective Start measures, as were the other three tasks.

From the content analysis Task 4 Cog had marginally more mistakes reported. However the researcher does note that many participants reported that they found the cognitive loading added far greater difficulty to the task. This is shown by the more consistent level of reported demand compared to other tasks.

4.10.3 Purpose – Identify what types of errors are made when using the controller and the possible cause for these.

Errors are identified in results as Slips, Lapses and Mistakes. These errors were not directly measured but other quantified measures, namely self reporting and performance times, were used to indicate what types of errors are likely to be occurring and to what extent. Groups <35 and 50+ are examined for this purpose as DNC group did not complete all experiments and many of the participants in DNC experienced unique errors due to higher ICF levels.

From the combined results it is clear that:

- Mistakes and lapses are the prominent error types
- The 50+ group had more reported and measured errors
- Errors are indicated by slower performance and greater reported task loading

The content analysis performed on the data obtained from the close out interview data was used to gauge the participants' awareness of type and number of errors made. Changes in Performance Time were used to indicate reductions in error, as an improved performance time would correlate directly with fewer errors. Self reported demand and performance would also correlate directly with the number and type of errors. The change in reported demand and performance would indicate the participant experiencing fewer errors.

Many participants were observed to become caught in 'mistake and lapse error loops'. That is they were wrongly recalling what they had done previously (lapse error) which created confusion. Because of this they began to make new mistakes which would lead to lapse errors in the next task as the mistakes prevented them from forming accurate metal models of the system.

Content Analysis

The content analysis indicated that more than half of participants were aware of mistakes, with 50+ reporting more than <35. Few participants could identify on which tasks and attempts mistakes occurred but those who did reported mistakes occurring on Attempt 1 on a range of tasks.

Table 4.109: Error comparison between <35 and 50+

Percent of participants with confirmed or possible errors on one or more tasks		
	<35	50+
Mistake	43%	75%
Lapse	58%	64%
Average	50.5%	69.5%

Lapses were reported as being more common than Mistakes, and with more uncertainty about whether they did occur. For the few participants who identified when lapses occurred they were reported solely on Task 4 Cog. A percentage comparison of reported errors can be seen in Table 4.19.

Slips were reported as being minor, only 2 of 26 participants reported occurrences and a further 2 reported maybes.

Performance Time

Group 50+ had slower Performance Times, reported greater levels of TLX demand and lower levels of TLX Performance (see Linear Mixed Effects Model, Table 4.10). They also required greater assistance when completing the tasks, indicated a lower confidence with using new technology and gave a lower SUS compared to <35 (see Independent T-test Table 4.6). Comparing this with the results of the content analysis there is strong evidence that errors correlate with slower performance and higher self reported levels of task load. With more attempts Performance Time improves and both reported task load and reported errors decrease.

4.10.4 Additional Outcomes - Observations: Reliability of self-reported measures; Effects of Age on Performance; Influences on Self-Reported Usability

Reliability of self-reported measures

This study used a mix of direct measurement, recording Performance Time and indirect measurement, self-reported task loads, system usability and interviewing. This allowed for a comparison of these different methods to gauge reliability of self-reporting.

If measurement of Performance Time is considered to be the 'purest' representation of performance on a specific task then self-reported task load and system usability as a whole, follow.

Group DNC gave a mean SUS of 2.53 out of 5 despite these participants being unable to successfully complete the experiment.

50+ gave a mean SUS of 2.21 despite reporting high levels of dislike and error in the close out interview. They also had a mean Performance Time and Task Load measures for the beginning of each task, except for Task 2 Light. It is likely that these times would be considered unacceptable in industry, for example some participants taking 7 minutes to self-teach how to turn the clock on and off.

Group <35 likely gave the most realistic SUS of 3.01.

Task Load tended to align with actual performance for <35 and 50+ with slower Performance Time resulting in greater task loads and vice versa throughout all tasks. This did not apply to DNC whose members had a worse actual performance than the other groups but generally reported task loads between <35 and 50+.

Effects of Age on Performance

This study focuses on effects of impairment, represented by ICF. However the participants recruited and the groups they fell into have allowed for the effect of age on performance to be observed. Specifically less than 35 years of age compared to over 50 years where neither groups has an ICF value over 1 and therefore no statistically significant difference in ICF.

It is clear that there is a difference in performance, 50+ having a slower Performance Time compared to <35 with a greater range (see Figures 4.5 to 4.8, Table 4.6 and Table 4.10).

However by the fourth Attempt on each task the difference in Performance Time between groups reduces to no more than 45 seconds (see Table 4.10).

The majority of task loads and usability are also statistically significantly different between <35 and 50+, with 50+ reporting greater task loading and a less usable system. Group 50+ also reported consistently greater Mental Demand and Frustration at the end of all four tasks, and greater Temporal Demand at the Start and End of all but Task 4 Cog.

Influences on SUS

The reported usability by participants appeared to be most influenced by Task 3 Clock.

Task 3 Clock can be considered the most difficult as it has the slowest Performance Time (Table 4.10), the highest cognitive task loading of the first three tasks and is comparable with Task 4 (Figures 4.15 and 4.16).

From the Multiple Regression assessment two statistical models had significant effects on SUS; these were the combined TLX Start Demand of the four tasks and the combined Attempt 1 performance time of the four tasks. Other statistical models were tested and were found not to be significant. In these statistical models the only attempt that had a statistically significantly effect on SUS was Attempt 1 Task 3 contributing 20.5% to the variance of SUS. Two demands were significant, TLX Start Demand Task 1 contributing 24% to variance of SUS and TLX Start Demand Task 3 contributing 31.1% to variance of SUS.

5.0 Task Process Model

5.1 Overview of the TPM

A model was developed based around the active change from one state to another due to an actor who consciously or sub-consciously is driven by an abstract desired state. This concept is consistent with many of the internal models seen in the literature review (section 2.6), with the greatest influence from Miall & Wolpert (1996). This base of the model is consciously driven state change which captures the idea of goal driven actions.

The model has been named the Task Process Model (TPM) and is intended to represent a person approaching a situation to complete a task with an intended outcome. The basic concept of the TPM can be seen in Figure 5.1, a key aspect of the model is that it is designed for tasks they are trying to achieve that have a desired and predictable outcome (as opposed to an abstract and complex problem-solving task).

The TPM describes a person beginning a task from a Beginning State (BS) with the intentions of changing the BS into a Desired State (DS). The person will take action (User Action) to move from the BS to the DS which will results in a New State (NS). The User Actions do not occur in an isolated environment and can be affected by external influences. The NS is then compared to the DS to confirm if the DS was achieved, this Comparison and Review represents how people reflect on how effective their actions were.

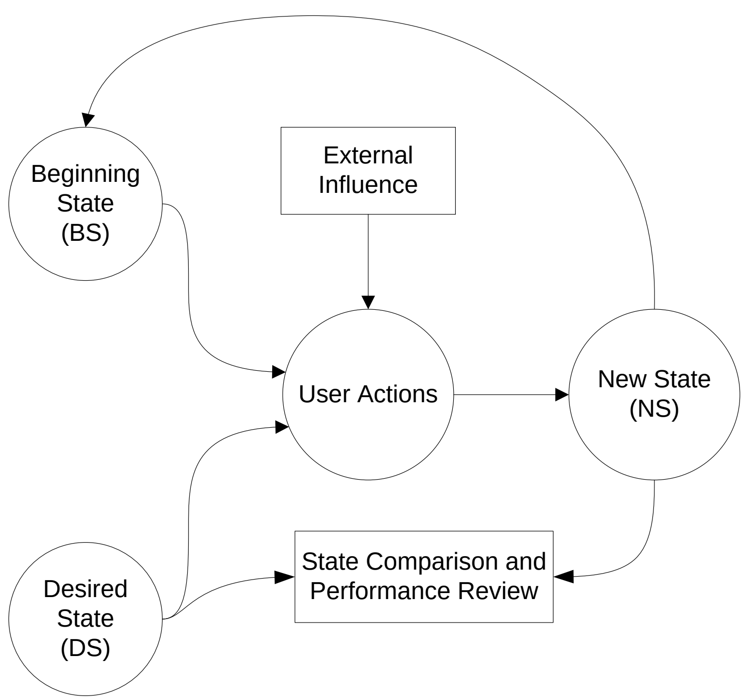


Figure 5.1: Basic Concept of TPM

The full TPM can be seen in Figure 5.2 which expands the Users Actions into a cognitive aspect and physical aspect, these are seen as Cognitive Recall and Assessment and Motor Action Output respectively. The External Influence component affecting the User Actions is also split into User's Innate Ability (such as skill, impairment, experience and age) and Cognitive Loading. The influence of error is also represented with Mistakes or Lapse errors affecting Cognitive Recall and Assessment and Slip error affecting Motor Action Output. Finally the State Comparison and Performance Review is linked with the Subjective Response representing the outcome of how a person perceives their ability at performing the task which then influences the perception of the New State and in the turn the next Beginning State.

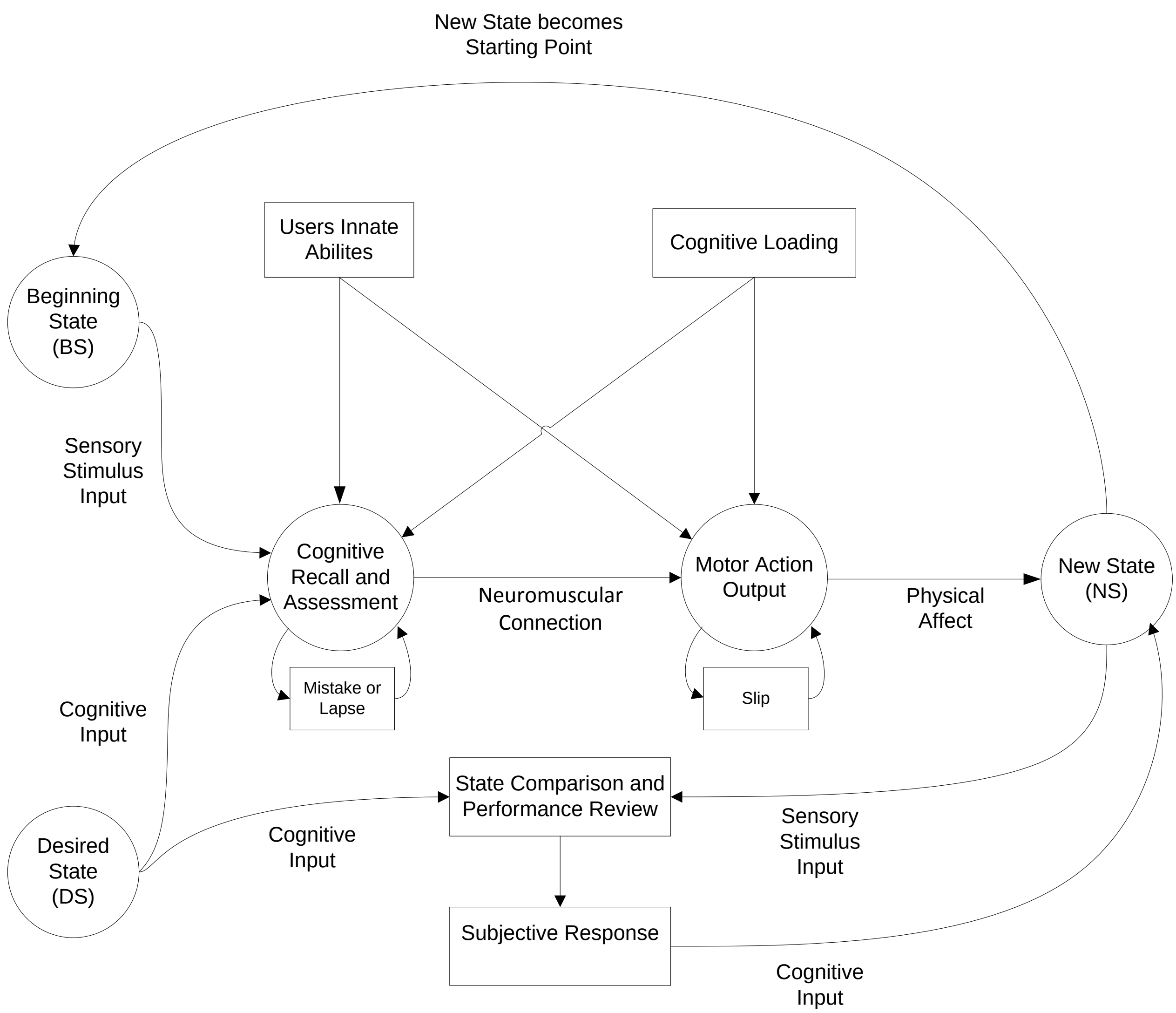


Figure 5.21: Task Performance Model

5.2 General Application of the TPM

Referring to the full TPM in Figure 5.2 the process of users performing a task can be described by the TPM through the following steps:

1. User approaches task from the Beginning State (BS) which includes the current situation and mind set of the user, and their surroundings. The user also has an intention of what they want to achieve, this is the Desired State (DS).

2. The difference between the BS and DS leads to a Cognitive Recall and Assessment by the user to determine the process and actions need to bring the BS to the DS. Here previous similar or identical tasks to the one at hand are recalled to plan the needed actions

Cognitive Recall and Assessment is influenced by the user's Innate Abilities and any cognitive loads on the user. The clearest effect on the user would be a decrease in speed completing the task, as was seen in the comparison of tasks performance times between groups <35 and 50+ (see Table 4.3).

3. During Cognitive Recall and Assessment errors of Mistakes and Lapses may occur. A Mistake or Lapse is caused by cognitive processing and will in turn influence it, most likely slowing the process. This is more so with a Mistake if the user is aware they are likely making an error.

The probability of a mistake or lapse occurring may be altered by the User's Innate Ability and Cognitive Loading.

4. A Motor Action Output is performed after a plan to act is decided upon through Cognitive Recall. Here Slip errors may occur which alter the intended performed action

The Motor Action Output is affected by the Users Innate Ability and Cognitive Load, potentially increasing likelihood of Slips and altering performance speed.

5. Once motor actions are performed by the user on the BS a New State (NS) is created. The NS becomes the BS for the next set of planning and action. The NS is also mentally compared by the user to the DS to determine if their goal was achieved.

6. The user will also review their own performance in addition to comparing the DS and NS. This serves two purposes. First the performed task and its success are stored as experience

for future recall. Secondly the user compares how easy they thought the task would be to how easy they found it. Any difference may be due to their performance or the challenge of the task being different to what they expected.

7. The State Comparison and Performance Review directly influences the user's subjective response to the system; how usable they found it, whether they enjoyed it or not and what their lasting impression will be. The subjective response of the user then influences the NS which in turn affects the next BS and the Cognitive Recall and Assessment.

5.3 TPM Application to Study Results

The TPM can be applied to a very specific action such as pushing a single button as well as larger more complex tasks. Here the TPM is applied to participant's interaction with the wheelchair controller which would be considered a larger and more complex task, seen in Figure 5.3 (over page).

The applied TPM in figure 5.3 incorporates the empirical results from this study to refine the model for application to this particular situation. The process is similar to that described in section 5.2 above but incorporating key results of this study to give specific application for use when designing for the wheel chair controller.

The following refers to the numbered grey boxes in Figure 4.25. The two main factors identified which affected use interaction with the controller were ICF and Age; other relevant factors which did not identify users are noted as 'General'.

1. Users Innate Abilities

ICF: Participants have 1.6 times greater risk of failure here through mistakes and incomprehension for each ICF level they gained (based on the survival analysis). Their ability to learn and acquire new skills having the greatest impact on success.

Age: Participants have a 64% slower performance time with this difference applying predominately to the first attempt (based on the Linear Mixed Effects Model)

2. Cognitive loading

General: An auditory response task creating cognitive loading prolongs performance time by an average of 17 seconds for all users once practice effects have occurred. This affects both speed of motor output and cognitive recall and assessment, see Table 4.9.

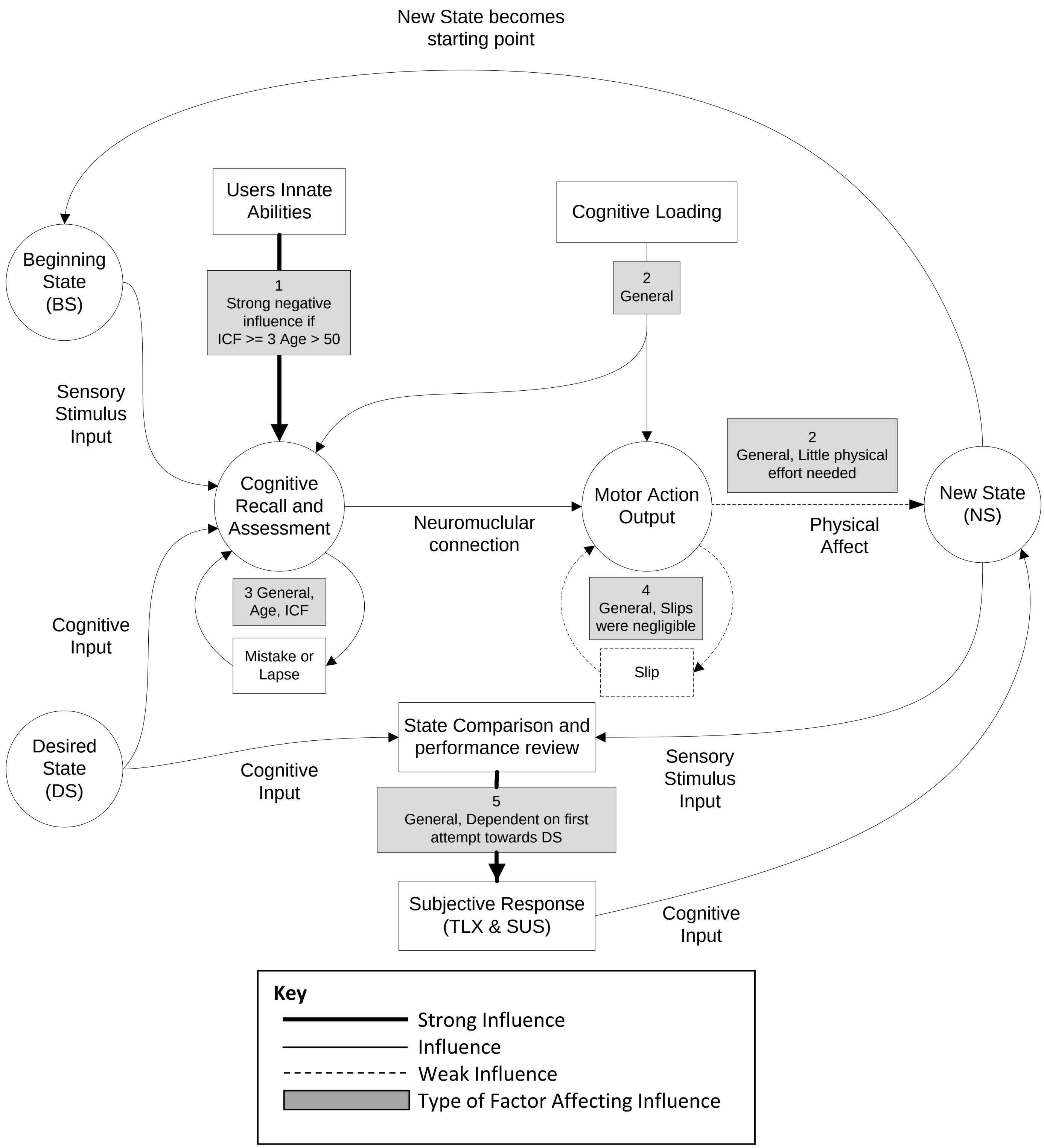


Figure 5.3: Task Performance Model applied to study

3. Mistake and Lapse Errors

(Based on content analysis)

General: 57.6% of participants reported at least one mistake mostly happening on the first attempt. 76.9% of participants reported at least one lapse, presumably occurring on later attempts.

Age: The 50+ group had 32% more participants report mistakes and 8% more report lapses compared to the <35 group.

ICF: Participants with statistically significant ICF measures compared to the mean of this study reported 46% of the total feelings of confusion when operating the controller. It is assumed this would lead to greater mistakes and lapses.

4. Slip Errors

General: Slips were negligible

5. Subjective Response to System

General: The perceived usability of the system was approximately 50% dependent on how the participant reacted to the first attempt of the most difficult task (based on stepwise regression)

The perceived performance and difficultly of the first attempt at any task determines user feedback more than performance by approximately 60% (based on Multiple Regression)

Reported Usability is more generous than what actual performance indicates.

6. Physical Affect

General: Little physical effort is needed to operate controller. Few participants struggled with the joystick of buttons.

5.3 Relation to Reviewed Literature

To allow the TPM to be representative of how people interact with a computer system it is important that it includes the three aspects of HCI identified by the author (Functionality of the system, Users ability and Communication of the system model to the user), see section 2.2.

These were incorporated into the model in varying degrees. The user's ability was able to be directly addressed through the ICF measure while functionality of the system is broadly addressed by the number of errors made and how closely the desired state and new state compare. Communication of the system model to the user is represented by how well the user performs in the Cognitive Recall and Assessment phase which is best determined by the measured performance of the user (how their desired state compares the new state) and their subjective response.

By including components that can be described as quantitative measures the TPM allows not only concepts of psychology and HCI to be represented but also provision of feedback. The TPM can point towards what should be measured to gauge the user experience as well as provide a means to apply that information directly to the model allowing it to become customized for the specific product as shown in section 5.2.

For further discussion of the TPM see sections 6.1.6 and 7.1.4.

6.0 Discussion

In this section the achievements of this study are discussed, including the practical benefits for industry and the novel intellectual contributions from the study.

This study had four key purposes identified in the Introduction, these were:

1. Identify the effects of impairments on a user's measured performance, and self-reported performance and usability experience.
2. Determine how cognitive loading affects a user's measured and reported performance.
3. Identify what types of errors are made when using the controller and the possible cause for these.
4. Create a model that can be used to explain the influence of the measured variables.

In this discussion these intentions are addressed along with the limitations of the study, implications for practitioners and potential future research.

6.1 Outcomes

This study has had several outcomes. First is the practical benefit of a usability assessment on the wheelchair controller for the industry partner.

Four research contributions were made in the areas of effect of impairment on measured performance, effect of cognitive loading on self-reported performance, effect of age on performance and the differences between measured and reported performance.

Finally a conceptual model, the Task Process Model, was developed and applied to the results of this study.

6.1.1 Usability assessment of wheel chair controller

The usability assessment of the wheelchair controller benefits the industry partner in several ways. First it provides a direct metric for performance of the controller by indicating the rough learning curve of the various tasks. Secondly, usability of the controller was reported through the TLX and SUS. The industry partner has been able to use the information provided for broad product design planning as well as specific redesign of the controller.

Problems in Software Over Hardware

The broader concept of ergonomics compared to interface design is the most beneficial for the industry partner in terms of practical design. The results of this study clearly showed for this controller that any usability issues the participants had were almost exclusively due to the digital interface rather than the physical ergonomics of the controller.

High Impairment Drastically Reduces the Controller Usefulness

Through the quantifiable metric of the ICF clear results have been provided for the industry partner. Most notable is the inability to complete the three tasks if a participant's ICF is 3 or more. Paradoxically these are the people who are most likely to use the controller.

Recommendations for Industry Partner

The key take away points for the industry partner are as follow:

- Base the menu structure on how the user uses the controller rather than similar feature grouping. For example rather than having a menu screen with 'Lights, Indicators and Hazard Lights' group features for scenarios the user encounters such as an outdoor use option with higher speed, lights on, chair lowered and slightly reclined.
- Consider removing unused features, such as screen background colour choice
- Redesign icons that can be confused, for example lights, background colour and brightness
- Redesign menu navigation
- Reduce number of menu screens
- Provide a reference and return feature in the menu navigation

6.1.2 Effects of impartment on usability and performance

In the literature there is significant research on usability design for people with various forms of impairment, disability and the elderly in both physical products and software design (Crews & Zavotka, 2006, 2006; Federici et al., 2005; Jaeger, 2009; Keates, Clarkson, Harrison, & Robinson, 2000; Newell & Gregor, 2002; Jakob Nielsen, 1995; Palmeri, 2006; Petrie, Hamilton, King, & Pavan, 2006; Pullin, 2009; Rose, Brooks, & Attree, 2002; Rowan, Gregor, Sloan, & Booth, 2000; Rowan et al., 2000; Shneiderman, 2000; Theofanos & Redish, 2003; Waller, Langdon, & Clarkson, 2010).

However, as mentioned in the literature review, there appears to be little to directly quantify impairment and specifically relate it to measured usability. Using the ICF to quantify impairment is a novel contribution. None of the studies and texts referenced use the ICF with product design despite directly discussing disability impacting on usability. It is not clear why the ICF is not more widely used in usability analysis applications. A possible explanation is that disability is often addressed at the beginning of the design process and in a specific and absolute way rather than in a general and discrete sense as provided by the ICF. For example products are designed for a people with poor eyesight rather than considering a more complex mix of impairments such as mildly poor eyesight combined with poor motor control and a resistance to acquiring new skills.

From the results given by the independent T-tests (sections 4.2 and 4.3) there is a clear correlation between greater impairment levels gauged by the ICF, and worse performances and a greater chance of not completing the outlined tasks. The impairment measure of Acquiring New Skills was the most influential contributor to poor performance. This is understandable as the experiment was based on a person's ability to learn how to operate a new device predominantly by self-teaching.

From the TLX, interview results, and observation by the researcher it appears that having a reduced ability to acquire new skills affects performance in a more indirect way. Most participants appear likely to be capable of learning how to use the controller. But when having to self-learn in a possibly pressured environment such as the experiment, levels of stress increase and disrupt the cognitive learning and recall process. This likely creates a negative feedback loop which reduces performance.

The TLX shows this by indicating that the 50+ group who had worse performance compared to the <35 group reported greater temporal demand (feelings of time pressure or being rushed) despite no time limit being placed on participants. Likewise the content analysis of the interviews showed high levels of confusion for all groups, suggesting pressured learning.

It could be argued that the experiment was biased to cause results of poor performance and usability by having participants self-learn. However the <35 group have exceptional performance proving that the experiment was not overtly biased, if it was bias at all. The intention was to test the usability of the controller in a somewhat realistic environment where help from a trained operator or user manual is not available.

Ultimately this study has shown that there is a strong link between reduced performance to the point of inability and greater impairment as measured by the ICF, at least in terms of operating this specific wheelchair controller.

6.1.3 Effects of cognitive loading on usability and performance

An objective of this study was to determine how cognitive loading affects a user's measured and reported performance. In review the experiment was not ideally designed to identify the specific effects of cognitive loading. Cognitive loading was applied to a combined version of the three tasks after the participants had gained experience using the controller, because of this there is no pure control for comparison.

Ideally, with more participants, the three groups would have been split into two sub groups. One of these would have cognitive loading applied and one without, to gauge a more direct and accurate comparison of the effect of distraction on measured and reported performance.

It is known that cognitive loading has a negative effect on a person's ability to problem solve (Sweller, 1988) and acquire knowledge (Brunken, Plass, & Leutner, 2003; Paas, 2003). In this study extraneous (external) cognitive loading was applied as a word association task.

The results suggest that cognitive loading had an effect on both measured performance and reported performance. The effects on group DNC were not measured as none of the participants got to Task 4 Cog to complete it so these conclusions are based on groups <35 and 50+.

The effects of cognitive loading were not particularly severe. Although performance time was slowed it was not a substantial loss and the greatest affect was that participants appeared to report higher frustration during Task 4 Cog. This suggests that getting experience and practice with the controller through the first three tasks effectively mitigated most of the adverse effects caused by the word association task.

From observing the participants it was clear that those with worse performance experienced greater frustration which created a positive feedback loop leading to worse performance followed by greater frustration and so on. This phenomenon can be seen in Figure 4.11 and 4.12 where several participants in group 50+ had slower performance times in later task attempts which go against the normal trend. Similar to this is the effect of experiment fatigue which is addressed in section 6.3.2.

The greater task load is likely attributed to a few specific aspects of the tasks, most notably the Clock Task as it provided poor feedback to the users. Becoming distracted at the wrong time would result in confusion of where one was in the process. This affected performance but had a greater effect on emotional response.

In terms of design application this means that a device such as the wheelchair controller tested may be confusing to learn but simple once understood and is able to be used effectively despite distraction. However the users are likely to still experience frustration, fatigue and other aspects of task loading.

6.1.4 Effects of Age on usability and performance

From the comparison of groups <35 and 50+ there is a clear result that age had an impact on measured performance and reported usability. The 50+ group had a slower learning curve that is likely to be solely due to older age. The initial time taken to complete tasks was slower for the 50+ group compared to <35 but by the final attempts on all tasks both groups had near identical performance. Despite this final similar performance, the reported usability through the SUS and the task load through the TLX was greater for the 50+ group.

It is likely that the younger age group could learn to use the controller more quickly as a younger person's brain is simply faster at learning new skills (Burke, 2006; Casey, Tottenham, Liston, & Durston, 2005) and younger people are more experienced with technology and managing multiple tasks simultaneously making leaning a new electronic device easier (Carrier, Cheever, Rosen, Benitez, & Chang, 2009; Judd & Kennedy, 2011; Shors, Anderson, Curlik II, & Nokia, 2012).

Although there is a large amount of literature on how learning ability changes with age the results here are novel as it also involves self-reported usability, which captures the emotional response of the participants.

6.1.5 Perceived usability and reliability of self-report

This study has clearly shown that greater impairment as measured by the ICF has a direct negative effect on performance. The reported usability measured by the TLX and SUS does not have a clear link with impairment, as group DNC who had the highest levels of ICF measured impairment and worst performance did not give the lowest usability scores. However looking at each of the three groups separately it is clear that with worse performance there is a reported lower usability compared to other participants in the group who had better performance.

This phenomenon is most likely attributed to response biases, specifically social desirability bias, experimenter's bias and demand characteristics bias. These biases are known to have a potentially significant effect, particularly the social desirability bias (Arnold & Feldman, 1981; Chung & Monroe, 2003; Fisher & Katz, 2000; A. Furnham, 1986; Nederhof, 1985; Paulhus, 1991; Weber & Cook, 1972).

Social Desirability Bias

The social desirability bias states that participants tend to deny undesirable traits, and ascribe to traits that are socially desirable (Paulhus, 1991). Participants with greater impairments who could not use the controller claimed that it had reasonable levels of usability; this may be explained by these participants simply not wanting to admit their inability or insult the design of the controller. As a participant in the Did Not Complete (DNC) group who gave a relatively high usability score said "I couldn't use it that well but I'm sure it's good and other people can use it."

Demand Characteristics and Experimenter's Bias

The demand characteristics bias states that participants alter their response or behaviour simply because they are part of an experiment (Orne, 1962). Similarly the experimenter's bias states that participants will tend to produce results that they believe the experimenter is looking for (Orne, 1962). In this study it is likely that participants when reporting on the TLX gave responses more in line with what they thought the researcher was expecting, possibly reporting greater loading or less loading over time.

Response Bias

Response bias is a general term for a wide range of cognitive biases that influence the responses of participants away from an accurate or truthful response. These biases are most prevalent in the types of studies and research that involve participant self-report, such as structured interviews or surveys (Furnham, 1986).

The literature does not appear to address if people with impairments are more or less likely to fall into response bias. Looking at the results of the SUS this study shows that performance aligns with reported usability according to the <35 and 50+ groups. Looking at the DNC group this pattern does not continue. This suggests that people with impairments may have a greater tendency to fall into some form of response bias, that is provide what they believe are the desired outcomes for the experiment.

This greater tendency for the DNC group to fall for response bias may be due to several reasons. Firstly the perceived authority of the researcher and the testing of an already made product may cause participants to give positive feedback. Secondly the participants may not want to appear inept due to the their impairment and age causing them to give positive feedback. Finally the design of the overall design of the experiment and specific phrasing of the questionnaires may bias the participants into a positive response.

Practical implications of this are discussed later in section 6.2.2.

6.1.6 Task process model

In this section the Task Process Model (TPM) is discussed in full, including application in practice, limitation and future work.

The Task Process Model (TPM) is a novel intellectual contribution from this study, presented in Section 4.10. It is based on various human behaviour and information processing models examined in the Literature Review Section 2.6.

Tasks Applicable to the TPM

The TPM is intended to describe the basic process of how a person may approach a general task in most situations. The term Task is intentionally broad to encompass any situation a person is in where they ultimately perform a physical action based on cognitive processing to create a change in their environment.

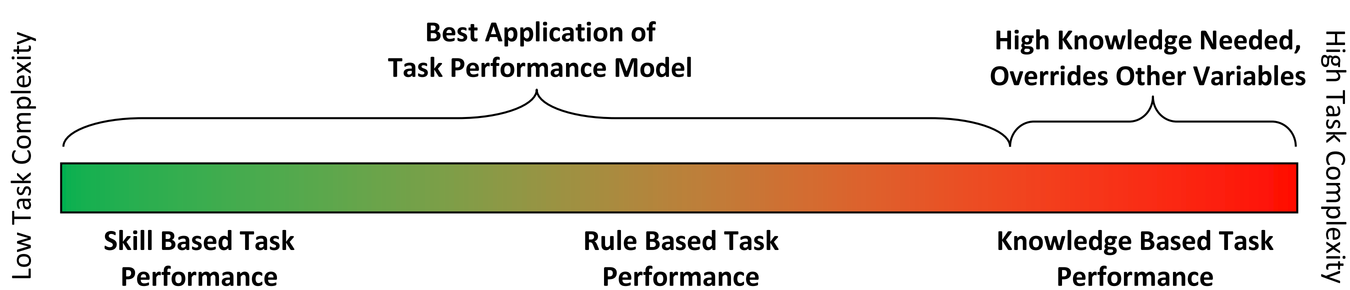


Figure 6.1: Complexity spectrum of task performance (Image by R Horne)

As seen in Figure 5.1 the tasks that TPM are best applied to are in the skill and rule based performance levels as outlined by Rasmussen (1982) and Reason (1995), overviewed in section 2.3.2. The TPM can still be applied to knowledge based performance task but highly cognitive tasks with little or no physical impact on the environment such as reading would not be well represented by the TPM.

Tasks in the skill and rule based performance levels may range from being simple, such as brushing ones teeth, to the complex such as trying to program a delayed recording on a television for the first time. The difference between these complexity levels is the amount of effort spent in the Cognitive Recall and Assessment step compared to the Motor Action Output Step.

Bushing one's teeth takes very little effort to recall how to pick up a brush, apply tooth paste, etc and put some reasonable physical effort to move the bush up to and around one's mouth.

In comparison, programming a modern television to record for the first time would require little physical effort, only pressing buttons on a remote control which is a well-practised action. However the cognitive assessment and recall element is substantially greater than brushing one's teeth. Previous attempts at programming other televisions, DVD players and VCR machines may be recalled and the on screen instructions, images and prompts would be carefully assessed as a mental model is built.

Tasks which fall too far into knowledge based performance become complex tasks. At this performance level other less cognitively demanding variables, for example external noise, which affect how the user performs the task become insignificant, as seen in Figure 6.1. For example attempting a new and complex maths problem is physically easy, requiring nothing more than basic actions such typing, writing and reading. However there is high demand for recalling previous experiences and cognitive processing and the affects of cognitive load become unpredictable. In addition solving the problem is not a simple step by step process with many possible means to answer it creating a high chance of mistake errors. This prevents the TPM from being applicable as there is a large difference between the beginning state and desired state and method to move between them is unpredictable.

Application of the TPM

The TPM can be applied to a task from a macro perspective or micro perspective, seeing a task as a single goal or breaking that single goal into smaller and smaller sub-goals or tasks respectively. This ties in closely with the concept of the Task Analysis (B. Kirwan & Ainsworth, 1992; Militello & Hoffman, 2008). For example the task of brushing your teeth could be explained with one action point 'brush teeth', it could be broken down into five steps, or into a further into 22 step as seen over the page in Table 6.1.

Each of the action points seen in the different approaches in Table 6.1 could have the TPM applied to it. The level of task break down is at the designer's discretion. A particular sub-action may be prone to error and the TPM may provide insight into how the user is approaching the task and what may be causing the errors.

The TPM is intended as a tool to give insight to designers with user centered design focus. It allows insight into the psychology of the user as they approach a task and is able to mix with existing user centered and user experience design methods such as Hierarchical Task Analysis, Cognitive Task Analysis or as a basis for Usability Testing and Interviews.

As shown in Results Section 4.10.2 the TPM can be validated and adapted to be more specific for purpose through empirical evidence, as it was done for the results of this study. This validation can be done in a flexible manner with a select few components of the TPM being tested or all of them tested, allowing for quantified statistical influence between components to be found, similar to that shown in the results section.

Table 6.1: Comparison of Task Analysis complexity

Task Analysis of Brushing Teeth		
Broad Action	Simple Break Down	Detailed Break Down
1. Brush Teeth	1. Put toothpaste on toothbrush 2. Bring brush to mouth 3. Scrub teeth with brush 4. Spit 5. Rinse Brush	1. Prepare Brush 1.a Pick up the tooth brush 1.b Wet the brush 1.c Take the cap off the tube 1.d Put paste on the brush 2. Clean Teeth 2.a Brush the outside of the bottom row of teeth 2.b Brush the outside of the top row of teeth 2.c Brush the biting surface of the top row of teeth 2.d Brush the biting surface of the bottom row of teeth 2.e Brush the inside surface of the bottom row of teeth 2.f Brush the inside surface of the top row of teeth 2.g Spit 3. Clean Up 3.a Rinse the brush 3.b Replace the brush in the holder 3.c Grasp cup 3.d Fill cup with water 3.e Rinse teeth with water 3.f Spit 3.g Replace cup in holder 3.h Wipe mouth 3.i Screw cap back on tube 3.j Return tube 3.k Return tooth brush

Limitations of the TPM

The TPM has several limitations. It is a relatively simplistic representation of how a person approaches a task, particularly as it sums all cognitive and physical action in two components. Although the point of the model is to simplify these complexities and highlight what affects them the model should not be taken as a true and complete information processing model.

As a design tool the TPM is still in its infancy and requires practical application for validation. However for those professionals involved in product design (such as engineers, industrial designers or software developers) who often have little knowledge of psychology, the TPM can provide insight.

A key limitation of the TPM is that it does not address aspects of usability of the device or system being tested. That is, it has no distinct components in the model to account for affordances, clarity of instructions or intended emotive effects of design. This was in part intentional as it keeps the focus on the user. But mainly was not done as considering these additional variables was too complex for the scope of this study.

The TPM also does not quantify the limits of working memory with recommended levels of cognitive loading, user's innate ability, target error rates or subjective response. This was done to keep the model simple and general.

Future work for the TPM should include practical application. Full validation for a number of devices or systems with industry-based teams would begin to show the usefulness of the TPM as a practical tool and to identify its faults.

An application of the TPM could be developed with various experimental and analysis methods more clearly outlined. This would reduce barriers to its use and allow for easier testing.

6.2 Implications for practitioners

In this section practical application of the results are discussed. Two forms of practice are covered, designing for disability and general design.

6.2.2 Impairment and Disability Focused Designs

The following discusses how product designs aimed at users with impairments or disabilities can benefit from what was learnt in this study.

Use the ICF measure as a review tool

The ICF is a more general measure of disability and impairment and so would be poorly used in a specific design process. Knowing a person had a level of 3 for Acquiring New Skills on the ICF would probably not be exact enough to aid specific designs.

This study shows that the ICF provides a way to easily compare a sample of people with diverse abilities all using a single product. Combing the ICF with a general way to measure usability (such as the TLX and SUS) allows usability with reference to disability to be applied in universal way.

A critique of the combined metrics used in this study is that they will likely be most useful reviewing a product design. Beginning a product design with the goal of having users with a measured ICF of 5 report a high level of usability (SUS of 5) and low task load (TLX of 1) is unlikely to help the engineers in the design process other than as a broad goal. However, in review these numbers will have far more context as they can be associated with the current design and used as a standard.

Beginning the design process with ICF values as targets can still be done but there would need to be a well-established real world understanding of the ICF measure.

Impairment involves many variables

Two important findings from this study were that age caused a significant performance gap and that different forms of impairment had varying and unpredictable effects on performance.

The 50+ age group showed a considerably steeper learning curve and reported higher levels of task loading and lower levels of perceived usability. This was likely due to the generational difference of being less adept at learning new technology on a consistent basis (Bennett, Maton, & Kervin, 2008; Mcmillan & Morrison, 2006; Selwyn, 2009; Tapscott, 2008) and slowing of cognitive processing that comes with age (Horn & Cattell, 1967; Salthouse, 2009; Salthouse, Atkinson, & Berish, 2003; Volkow, Gur, Gene-Jack, Fowler, & al, 1998).

An effect of age on performance was identified as a lack of confidence resulting in some participants persevering with the tasks only because it was a formal study. A lack of confidence with new technology combined with a reduced ability to learn and/or operate systems and devices can create large barriers in users that designers need to overcome. This was seen in the study with the DNC group who were both of an older age and had various levels of impairment.

Comparing the results of the DNC group to an untested, hypothetical group of young people with impairment; it is expected that this group would be confident and open to learning new technology but have reduced ability to operate the controller. Designing for the group would have different constraints compared to that described in the previous paragraph.

This study highlighted that a person’s impairment is only a single aspect of how they would approach and interact with technology. Their motivation, experience and other variables are all important factors that should be considered in design, for example when creating personas as a design tool.

6.2.3 General Product Design

The following are lessons learnt from the study that can be applied by design practitioners.

Be wary of subjective measures

It was seen from this study that subjective measures of usability in the form of the SUS and TLX do not strongly represent actual performance. In particular seen in group 50+ was that a poor first attempt will have a lasting detrimental effect to perceived usability despite an improvement in performance. In general most participants were likely affected by response bias giving overly positive scores for usability compared to the interviews which were generally negative as seen by the results of content analysis.

It is recommended that any subjective quantitative measures (such as the SUS and TLX) be backed up with observed quantitative measures (such as timed performance and tracked actions), and in-depth qualitative measures (such as user interviews) and observation by the researchers of the participants operating the device of system.

Basing how usable a product is on a point based subjective questionnaires is a way to gather numerical data quickly that can be used to clearly communicate results through graphs and statistics. However designers should understand the weaknesses of these tools and supplement them accordingly.

Errors types are not clear cut

A goal of this study was to identify the types of errors made when using the wheelchair controller. This purpose was partially met but not to the ideal standard. While the reviewed literature was clear and rich with errors types, what causes them and the underlying psychology, the means to practically identify errors was less clear.

In this study the interview content analysis was the main tool for identifying error types with the quantity of errors merely estimated from performance results.

It is clear that there were a large amount of errors made by the participants who reported and had measured poor performances. These errors were not slips and it is likely that most errors were rule based and occurred in the 2nd, 3rd and 4th attempts followed by knowledge based errors which occurred in the 1st and 2nd attempts. It is hypothesised that most errors were likely caused by a mismatch of the system model and the participants' mental model, and based around menu navigation and unclear icons.

It is not clear the exact number of errors made by each participant, when and where these errors occurred, what their exact cause was and how to mitigate them. This could have been determined with a different experimental approach such as outlining the steps to complete a task, recording the steps taken by participants and then comparing the two followed by questioning the participants about their reasoning for their chosen path.

Identifying the paths taken by participants was attempted. However the equipment needed to accurately record button presses proved impractical to use within the time frame of the study and for the industry sponsor to assist with.

It is recommended that exact error types be measured for designers working to polish an already existing product. The results of this study have provided information that can be used to influence future design changes for the wheelchair controller. However identifying the exact errors would have little impact as new common errors will appear and current errors will likely be dealt with.

6.2.4 Specific Recommendations

Following are several direct recommendations for product designers who are designing for people with impairments.

- **Understand the Intended and Likely Users:** A key aspect of user experience designer is to begin with a thorough understanding of the possible users. Based on this study this is particularly important when designing for people with impairments as their needs, ability and personal goals will likely have a wider range than people who are able bodied. It is recommended that the type of user is clearly identified, particularly the mix of impairments and types of life style that users will have rather than designing for a specific impairment.
- **Conduct Rigorous User Testing:** As it is likely that people who are able bodied will be designing for people with impairments, user testing the product with the intended user is particularly important. It is highly likely there will be a difference in the designer's physical ability, cognitive ability and lifestyle compared to the impaired user.
- **Minimise Product Presence:** Products designed for aid are there to enable a person to overcome a difficulty and allow them to reach their intended goal, rather than the use of the product being the intended goal. For example a person may play a game on their smart phone with the goal of enjoying the game design and challenge, but a person makes a call

from their smart phone with the intention of talking to another person rather than appreciating how the call is dialled. Similarly aid devices should have as little presence as possible and absolutely avoid frustrating the user by taking up as little time to use, requiring as little cognitive effort and being as discrete as possible. This may results in removing features and physically changing the product.

6.3 Limitations in the work

The limitations of the study are discussed, including the experimental methodology, data analysis and general experimental design. The limitations of the Task Process Model and the limits of practical application were discussed in Sections 6.1.6 and 5.2 respectively.

6.3.1 Uncontrolled variables in Experimental Process

Throughout the experimental method there were several uncontrolled variables that would have had an impact on the data collected. These impacts were minimized where possible but the extent of their influence is relatively unknown. The following examines the uncontrolled variables with the greatest impact.

Participant Down-Time

During the tasks there was time when participants were not doing anything, such as between tasks or the moments before they were asked to begin. During this 'downtime' participants would have been able to cognitively process and plan how they would approach the task, concisely or sub-consciously. This downtime may have caused a greater discrepancy between participants of different abilities, because of this downtime was minimised by keeping the participants on task to prevent its impact. However due to the nature of the experiment it was impossible to prevent any down time completely. It is not thought that the results would have been significantly affected by downtime but it should still be considered in any future experimental work.

Uncontrolled cognitive loading

The fourth task used a word association task to cause cognitive loading. The word association task involved the researcher reading the words to the participants as opposed to having the words given to them through a more controlled method such as audio recordings. The rate that the words were read was adjusted at the researcher's discretion dependent on the response of the participant. This enabled the researcher to adjust the cognitive load. The level of cognitive loading was not monitored, for example by measuring time taken for a participant to respond to a word.

A more controlled method of applying cognitive load would need to be used if the effects of distraction were a larger component of the study. However as the effects of cognitive loading were being explored on a more general basis rather than in detail this less than ideal controlled application of cognitive loading was acceptable.

Applying cognitive loading onto three combined tasks at the end of the experiment gave an indication of the overall effect of cognitive loading. This did not allow for any insight into how cognitive loading affected the initial learning process, only on the recall and application of existing skills.

This study has shown that initial attempts have a large impact on perceived usability and therefore it can be hypothesized that reducing cognitive loading during the early learning process would improve a person's subjective and lasting response to a device or system. With the number of participants in this study it was impractical to better analyse the effects of cognitive loading on each of the tasks separately.

6.3.2 Experimental Design

The experimental design aspects that were seen to introduce weaknesses and limitations are discussed.

Measured Performance Time Error

Participant performance time was measured by the researcher reviewing video recordings. Although care was taken and a frame by frame examination was used for the start and end of each task, accuracy was ultimately dependent on human observation and the time accuracy of the video recording.

However the error introduced though this method is considered to be minimal and did not affect the ultimate results of this study. If a greater emphasis was put on refining exact performance times for design refinement a different method would have been used.

Priming

The participants may have been primed to give poor SUS scores as the two hardest tasks, Task 3 Clock and Task 4 Cog, were the last tasks done before filling out the SUS. This would have been hard to mitigate and any arrangement of the task may have in some way primed participants.

Randomising the order in which participants did the task would have mitigated this limitation but would have required a far greater number of participants.

Experiment Fatigue

For some participants (mostly the 50+ group) the experiment lasted up to two hours or more, longer than the expected and planned experiment time. From observation it was clear that the participants who experienced longer experiment times were relieved when it ended, were reporting higher levels of frustration and appeared to operate the controller in a less structured way as they moved through the tasks.

Fatigue affecting the participant's performance is likely the uncontrolled variable with the greatest impact while also being the hardest to mitigate. The extent of the effect of experimental fatigue is unknown and may have created a wider performance gap between the participant groups. Those participants who struggled with the controller more would have taken longer to complete the tasks spent more mental energy and therefore suffered more from fatigue further decreasing their performance in a cyclic fashion. However the effect of fatigue is applicable to real world scenarios so it would not be ideal if it was completely removed.

To mitigate this weakness in the experimental design fewer tasks could have been tested or breaks for participants could have been given.

Sample Bias

The participants recruited were not ideally randomized. They were drawn from with the University of Canterbury though email and poster request adverts, through a number of different disability care organizations or by word of mouth. This means of recruitment resulted in the three distinct groups, young people from the university without impairment, people with impairments found through the various care facilities and the older and non-impaired group through word of mouth. The greatest bias was that the young people recruited were predominately university students who can be assumed to have an above average intelligence and familiarity with technology.

The sample bias in this study was unavoidable, for practical reasons the number and range of participants were limited, particularly those with impairments. However the study used enough participants to find statistical significance and has provided results that are likely to apply generally to the population.

6.4 Implications for future research

In this section potential future research leading on from this study is discussed.

6.4.1 Broaden Range of Disability and Impairment

A goal of this study was to identify the effects of impairments a user has on performance. No specific impairment or disability was sought and participants were recruited from third parties preventing the researcher from having much control over which participants were recruited.

This study did well to identify that impairment does have a significant effect on measured and subjective performance and that the performance of a person with an impaired ability to learn and acquire a new skill were most affected.

It is recommended that future research use the six identified aspects of the ICF which influence a person's ability to interact with as a device as a guide to recruit participants with specific impairments. Participants should be sourced who exemplify one of those measures to allow for a better comparison between the measures. Comparing the identified impairments more directly will provide insight into which impairments and more importantly at what level of impairment usability begins to be severely affected.

6.4.2 Identify and Purposefully Cause Errors

It was determined that errors occurred where knowledge based performance turned into rule based performance as participants became familiar with the controller and tasks, while almost no slip errors occurred.

Identifying the causes of these errors was primarily done by qualitative analysis of interviews and observation. It is recommended that if future studies are applied to refine the design of a particular device or system emphasis should be put on classing errors in more detail and identifying their root cause.

This can be done by more in-depth interviews with participants examining each task, defining each task with a task analysis and comparing this to what participants actually did and use of think-aloud protocols (Boren & Ramey, 2000; Fonteyn, Kuipers, & Grobe, 1993; Jaspers, Steen, Bos, & Geenen, 2004).

6.4.3 Refine and Explore Cognitive Loading

In this study cognitive loading was applied through a word association task on the final task. This provided results showing that cognitive loading had an affect on performance in a broad sense.

It is recommended that in future studies identifying what type of cognitive loading has the greatest effect on performance, a variety of different cognitive loads are applied. For example visual response, conversational or memory loading.

It is also recommended that cognitive loading be made a primary experimental test variable where a control group with no cognitive loading is compared to a group that only experiences the experiment with cognitive loading. This would allow cognitive loading to be applied from the first task and the full effect on the learning curve would be determined. In this study the external cognitive load was only applied in the final task after practice had taken effect.

6.4.4 Long term studies to test for retained learning

This study focused on a single session with each participant. It is unknown if over time the participants retained what they learnt about operating the controller or if their opinions changed on the usability of the controller.

Some of the tasks tested would be used rarely, turning the lights on may only be used monthly and accessing the clock feature only twice a year for daylight saving. The controller has many other features which may only be used once or twice such as changing the screen brightness or choosing a white or black background. If a person needs to be taught how to operate a system an important measure of usability is if that person would need to be taught the same lesson again. Features which are somewhat counter intuitive and infrequently used should be made more intuitive as it cannot be assumed that people will rote learn how to use them.

Knowing the long term usability for this specific controller or any industry product is important for success. It is recommended that attempts be made to return to participants for follow up testing.

In addition the retesting of participants would allow an assessment of how quickly an initial reported negative response to the system is overcome.

7.0 Conclusion

7.1 Purposes

7.1.1 Identify the effects of impairments on a user's measured performance, and self-reported performance and usability experience.

It was found that the impairments, as measured by the ICF, correlated closely with a negative effect on performance. No participant with an ICF level of 3 or more on any combination of the measured ICF impairments was able to complete all four tasks in the experiment. As measured by the Cox Regression Analysis a participant with an ICF of 3 or more was 181 times more likely to not complete the experiment compared to a participant with an ICF of 2 or less.

The impairment of Ability to Acquire New Skills as described by ICF was found to be the most critical followed by the quality of Psychomotor Functions and Control of Simple Voluntary Movements. This was determined by the Independent T-test and Cox Regression Analysis.

Specific ICF measures of impairments in Acquiring Skills, Psychomotor Functions and Voluntary Movements respectively had the greatest effect on performance.

Physically the majority if not all participants were able to operate the controller. Difficulties occurred due to the discrepancies between the participants' cognitive ability and the complexity of the controller.

Reported performance for participants with higher impairments was comparable to other participants. Those with higher impairments had a slight tendency to be overly optimistic about their ability and the usability of the system but this was not a strong result.

It can be concluded that even mild impairments relating to mental functions can affect a person's ability to operate the wheelchair controller and likely any device of similar complexity. In addition to dealing with impairments any initial poor performance is compounded by the frustration of trying to understand something that may be considered easy to use. Potential fatigue from previous multiple unpleasant attempts can lead to total inability to use the system. This can make a device which may have only a small amount of confusing design features become impossible to use by some, as was shown with the wheelchair controller in this study.

7.1.2 Determine how cognitive loading affects a user's measured and reported performance.

Cognitive loading in the form of external distraction through word association appeared to have an effect on participants' Performance Time and a greater affect on Reported Performance as determined by the Linear Mixed Effects Model and Mean comparisons of the groups. The final task where cognitive loading was applied had the second slowest performance time for all groups, with the most difficult task having the slowest performance time. It was also found that older participants were more susceptible to cognitive loading affecting both the measured performance and their reported performance. It is important to note that none of the DNC group, that is the participants with impairments, made it to the final task to be subjected to cognitive loading.

It can be concluded that negative effects of cognitive loading on performing a practiced task are not as detrimental to performance as having to learn a difficult task for the first time even with experience using the device. More broadly, knowledge based performance is more demanding than skill and rule based performance combined with cognitive loading.

The implication is that learning to use a system and the difficulties of that system is demanding. Designing it to avoid or deal with cognitive loading is important. This could be done by including navigation steps that can be paused or moved back.

The application of cognitive loading in this study was not done with a control group so all participants benefited from practice before cognitive loading was applied in the last task. Future research would benefit from the use of a control group and experimenting with different forms of cognitive loading.

7.1.3 Identify what types of errors are made when using the controller and the possible cause for these.

Errors were only able to be determined for participants who completed the experiment, that is the group of less than 35 years (<35) and the over 50 years (50+) group.

The majority of errors that participants made were mistakes or lapses (poor recall of process), with almost no slips. This reflects that there was almost no problem for participants to physically operate the system compared to the cognitive struggle to understand the system model of the interface.

The average percent of participants with confirmed or possible errors on one or more tasks was 50.5% of <35 and 69.5% of 50+.

Some participants were observed to become caught in mistake and lapse error loops, that is they were wrongly recalling what they had done previously (a lapse error) so became confused and then began to make new mistakes.

7.1.4 Create a model that can be used to explain the influence of the measured variables.

The Task Process Model (TPM) was generated to outline the cognitive process of a user attempting a task. The model is based on a change in state of the environment due to the user's actions and driven by the user's desired state.

The TPM is intended to convey in a simple way the complex cognitive process of a user's thought process while making a simple decision and a physical action. The target users of the TPM are practitioners who do not have a background in psychology but would benefit from a basic understanding of its application in usability design.

7.2 Additional Outcomes

The strongest additional finding was that Age had a strong impact on both measured and reported performance relative to impairment. This can be clearly seen by the statistically significant differences between the group <35 and group 50+ who differed on almost all measures of measured performance, reported performance, reported usability and reported errors, with the younger age group being more adept at using the controller. This is likely due to the measured greater confidence with technology of the younger participants as well as younger minds being quicker at learning new skills and more exposed to different technologies.

It was also found that the participants tended to self report more positive performances than the measured results indicated. Group DNC gave a mean SUS (System Usability Score) of 2.53 out of 5 despite all of these participants being unable to successfully complete the experiment compared to groups <35 and 50+ who gave mean SUS scores 3.01 and 2.21 respectively. A similar phenomenon was found for reported errors and TLX (task loading). This is likely due to several cognitive biases, namely the Social Desirability Bias, Demand Characteristics and Experimenters Bias and Response Bias. Essentially participants did not want to admit their inability to operate that controller nor give a negative review of what other people had designed and built.

7.3 Implication for Product Designers

The key point for practitioners designing for people with impairments is to understand that age is almost as strong an influence on performance as impairment. In this study the difference in performance of the <35 and 50+ groups is solely due to the difference in age.

It is known that a younger mind is able to learn more quickly but both groups were capable of performing all tasks to a near equal performance level after practice. The most likely reason that the younger participants performed better is simply that they are used to having to learn how to operate computer interfaces quickly.

This leads onto the second key lessons for designers. The initial subjective response of participants to a system is lasting and not easily overridden. Despite <35 and 50+ having near identical performances by the end of each task, 50+ reported feeling greater task loads and that the system was less usable. This is likely due to their struggles with the system during early attempts.

Finally user research and usability testing with users is highly recommended when designing for people with impairments. Designers are unlikely to be able to relate to the users on many levels; namely physical ability, cognitive ability and lifestyle. This highlights the need to close the gap between the user and designer and to allow the user to dictate how functional and usable the product is.

Appendices

See following pages for the appendicies

A 1 Published Article – Engineering Teams Performance

Industry-based Team-Projects: Personality traits that influence success in engineering education

Horne, R.; Pons, D.J. [6]; Helton, W.S.

Abstract

There is need to better understand how *personality* affects performance at *team projects* engineering education. The purpose of this study was to explore the relationship between final year mechanical students' personality traits, and their performance both in a capstone industry-based team-project, and their general academic performance through GPA. The final year project involved students working in teams of four with a supervisor and industry client; it included an individual mark and team mark. Final year mechanical engineering students were surveyed. From a class of 104 undergraduate students 43 usable student results were gained. Personality was gauged using a ten item instrument measuring the super traits of the Five Factor Model. It was found that conscientiousness had a significant correlation of 0.307 with GPA but not with either project mark. GPA had a significant correlation with project marks and accounted for approximately 16% of variance within both project marks. Individual and team project marks had a significant and strong correlation of 0.829 with each other. Openness and Agreeableness influenced the individual mark, and students who experienced some mild anxiety (Neuroticism) regarding performance in the Final Year Project performed better. This paper provides a structured model of proposed causality, which links personality, teams, and performance on a capstone industry-based team-project. This is important, because it suggests implications whereby the efficacy of the professional graduate attribute may be enhanced. Implications for practitioners are given, limitations of the study and directions for future research are discussed.

[6] *Please address correspondence to Dr Dirk Pons, University of Canterbury, Christchurch, New Zealand, dirk.pons@canterbury.ac.nz.*

Keyword: personality; engineering; teams; project; graduate attributes; education; academic success; GPA; Big Five; capstone; undergraduate; a very brief measure of the big-five

Introduction

It is important for the effective functioning of any profession that its members are able to work effectively as individual agents, while simultaneously coordinating their activities with other team members to accomplish larger outcomes. Therefore teamwork is an important part of the graduate attributes for any profession. However the team is made up of individuals with their own skills and personal attributes which they bring to the situation. Therefore the antecedent of team behaviour is the individual behaviour. Here we are particularly interested in one set of parameters that define individual behaviour, namely *personality*. We are interested in how individual personality affects team experiences, particularly in the educational setting. The particular area under examination is engineering, where the profession specifically identifies *teamwork* as a required graduate attribute (IEM, 2009). The contribution of this paper is the application of personality analysis to engineering teams in the educational setting.

A (Very) Brief Review of The Big Five

There are many ways to categorise and represent personality, but over the years the Five Factor model has come to dominate, and is briefly reviewed below.

The earliest studies investigating personality as a means of personnel measurement and selection suffered as there was not a well accepted taxonomy to classify personality traits (Barrick & Mount, 1991). Papers were soon published which united concepts of personality that many psychologists were converging on (Digman, 1990; McCrae & Costa, 1990; Poropat, 2009). This foundational taxonomy was the Five Factor Personality Model also known as the Big Five Personality Traits; Openness to Experiences, Conscientiousness, Extraversion/Introversion, Agreeableness, and Neuroticism (OCEAN). These traits were derived from statistical factor analysis of personality descriptors (or the lexicon) and are believed to be independent dichotomous variables. Each of these traits are rated for an individual on a scale with a high rating indicating strong trait tendencies and low ratings indicating opposite trait tendencies. The five traits are described for a high rating by Barrick and Mount (1991) as follows:

- Openness to experience describes a person who is imaginative, cultured, curious, original, and artistically sensitive.

- Conscientiousness describes a person who is dependable, has a strong will to achieve and is able to work hard.
- Extraversion describes a person who is sociable, gregarious, assertive, talkative, and active.
- Agreeableness describes a person who is likable, good natured, co-operative and tolerant.
- Neuroticism describes a person who is anxious, depressed, angry, emotional, worried, and insecure (the opposite of this is emotional stability as is used in the personality instrument in this paper).

These traits are intended to be non-judgemental: the opposite side of the trait is not intended to have a pejorative meaning. However, the very way that these traits are commonly framed in the literature (the above descriptions being an example), puts a positivist spin on certain ends of the dichotomies, and defines the desirable personality. Specifically, how many people want to think of themselves as unimaginative, lazy, timid, bad-natured, and angry? These being the simple antonyms to the above, with the exception of Neuroticism which is already negative.[7] Subjects may therefore fake good, i.e. self-assess their personality in the way they *want* to be perceived. Consequently the wording of the personality test needs to be non-judgemental, and to avoid this risk there are several standard self-assessment instruments available.

Personality and Intellect

The extensive literature on the relationship between the big five personality traits and intellect has come to several conclusions, as well as mixed results.

+ Openness: In terms of general intelligence Openness to Experience is the strongest predictor of intellect with positive correlations (Ackerman & Heggestad, 1997; Austin et al., 2002; Brand, 1994; Chamorro-Premuzic, Moutafi, & Furnham, 2005; Moutafi, Furnham, & Crump, 2003; Rosander, Bäckström, & Stenberg, 2011)(Ackerman & Heggestad, 1997; Barrick & Mount, 1991; Digman, 1990).

[7] The extent to which current definitions of the Five Factors frame and bias the personality construct is evident in a simple thought experiment, which is to describe high OCEAN ratings from the other perspective. Thus O: need for novelty and personal indulgence; C: workaholic and driven; E: fear of social disapprobation; A: weak-willed; N: emotionally aware. Framed in that alterative way, the other poles of the dichotomies immediately become more attractive and rational. That is not a perspective that the popular literature takes, nor even in the research literature. We therefore make the point that the Big Five is too often framed in a biased way: described with a preferred axis for each dichotomy. There is a need to move the debate to a more balanced consideration. Specifically, the negative attributes of high ratings urgently need inclusion.

- Conscientiousness: The correlations between conscientiousness and intellect have been mixed, however the most recent studies have found negative correlations (Allik & Realo, 1997; Chamorro-Premuzic & Furnham, 2008; Demetriou, Kyriakides, & Avraamidou, 2003; Moutafi et al., 2003; Moutafi, Furnham, & Crump, 2006; Moutafi, Furnham, & Paltiel, 2004).

+- Extraversion: Extraversion has been found to both correlate positively (Ackerman & Heggestad, 1997; Austin et al., 2002; Chamorro-Premuzic & Furnham, 2008; A. Furnham, Moutafi, & Chamorro-Premuzic, 2005; A.-P. Furnham, 2004) and negatively with intellegence (A. Furnham, Forde, & Cotter, 1998; Moutafi et al., 2003) although neither with a high correlation. This is thought to be dependent on testing conditions (Moutafi et al., 2006)

o Agreeableness: Agreeableness has been found to have no significant correlations with intelligence (Ackerman & Heggestad, 1997; Austin et al., 2002; A. Furnham et al., 2005; A.-P. Furnham, 2004; Moutafi et al., 2003)

- Neuroticism: Neuroticism has been found to have a negative correlation with intelligence (Austin et al., 2002; A. Furnham et al., 2005; A.-P. Furnham, 2004; Moutafi et al., 2003)

The results of a meta-analysis by Ackerman & Heggestad (1997) are seen in Table 1, shown is the correlation of the Five Factor Personality Model with Intelligence.

Table 1: Correlation of the Five Factor Traits with intellect

Study	**Five Factor Traits**				
	Openness	**Conscientiousness**	**Extraversion**	**Agreeableness**	**Neuroticism**
Ackerman & Heggestad (1997)	0.33	0.02	0.08	0.01	NA

From this it is inferred that Intellect is partly related to Openness, but not to the other personality factors. Not everything that makes up a person's behaviour is considered *personality*, and the trait-based approaches to personality generally endeavour to avoid including intellect. Apparently that has not entirely been the case with the Five Factor model.

Indeed, Openness is statistically the weakest of the Five Factors, i.e. all the residual variance is joined in that final factor.

Personality and Academic Performance

There have been many studies examining personality and academic performance (De Feyter, Caers, Vigna, & Berings, 2012; Hazrati-Viari, Rad, & Torabi, 2012; Kappe & van der Flier, 2010). Meta-analyses comparing the Five Factor Personality Model and tertiary academic performance (Poropat, 2009; O'Connor & Paunonen, 2007) show that Conscientiousness has the strongest correlation for academic performance, see Table 2. This contrasts with Openness to Experience as the main predictor for intelligence (Table 1). From this it appears that success at academic studies has more to do with Conscientiousness than Intelligence. There have been mixed results correlating Openness to academic performance and it is suggested that other moderating variables may be involved (O'Connor & Paunonen, 2007).

Table 2: Correlation of the Five Factor Traits with academic performance

Study	Five Factor Traits				
	Openness	**Conscientiousness**	**Extraversion**	**Agreeableness**	**Neuroticism (reverse ES)**
Poropat (2009)	0.10	0.19	-0.01	0.07	-0.01
O'Connor & Paunonen (2007)	0.06	0.24	-0.05	0.06	-0.03

It has been hypothesised that more intelligent individuals can have lower Conscientiousness as they are able to rely on their knowledge and intellect to complete cognitive tasks (Moutafi et al., 2006, 2004). Conscientiousness is thought to be linked closely to motivation as well as other sub facets which are likely beneficial for success in academic assessments (Chamorro-Premuzic et al., 2005; O'Connor & Paunonen, 2007).

What Makes a Good Engineer?

From the literature a typecast professional mechanical engineer is described in the psychology literature as having high levels of conscientiousness with a strong personal work ethic and goal orientation (Futrell, 1985; Harrison, Tomblen, & Jackson, 1955; Kichuk & Wiesner, 1997; Kline & Lapham, 1992); reasonable agreeableness and extraversion, as well as good 'people skills' and functional interpersonal relationships (Futrell, 1985; Harrison et al., 1955; Kichuk &

Wiesner, 1997); they are emotionality stable (Harrison et al., 1955); an adequate and well rounded intelligence (Futrell, 1985; Kichuk & Wiesner, 1997); and preferably a positive 'can-do' attitude with practical skills (Futrell, 1985). Few personality tests have been given to professional engineers in recent years. However a study of 103 professional engineers conducted in the Netherlands using the Five Factor Personality measure found that engineers had lower agreeableness but higher extraversion, conscientiousness, emotional stability and autonomy than the national comparison group (Van Der Molen, Schmidt, & Kruisman, 2007). It was also concluded that engineers with higher level degrees were less conscientious than engineers with lower level degrees (Van Der Molen et al., 2007). A recent study showed that students likely to choose to follow a degree in the applied sciences are likely to have above average levels of conscientiousness and below average levels of extraversion (Balsamo, Lauriola, & Saggino, 2012).

Independent of this psychology perspective, the profession itself identifies several attributes that are expected in graduates, including solving complex problems, dealing with conflicting issues, using abstract thinking, applying an analytical approach, dealing with diverse groups of stakeholders, synthesis of information, apply reasoning informed by societal and cultural issues, communicate effectively, lead teams, and manage projects (IEM, 2009). None of these are personality attributes per se, though the requirement for abstract and analytical thinking is consistent with *conscientiousness.*

Whether or not student engineers learn the desirable attributes is an open question. There is sometimes a perception in industry that graduates are scientifically adept but have weak soft-skills (NEEP, 2010). Nominally they do have the required skills, since all engineering programmes are accredited and periodically checked by the professional body, which in the New Zealand jurisdiction is the Institute of Professional Engineers (IPENZ) (IPENZ, 2006). The traditional examination-based model of engineering education has been criticised, suggesting that students take a superficial approach to learning the material and focus on passing rather than deeper leaning (Ditcher, 2001). Others suggest that desirable behaviour is learnable through an educational program (Newport & Elms, 1997). The degree to which engineering graduates have skills that match employers needs has sometimes been investigated (Dickson & Grant, 2008; Wadhwa, 1981) and results show that 'Graduates are apparently not as ill-prepared for the workplace as anecdotal comments from employers would suggest' (Banik, 2008).

One required professional attribute that is specifically relevant to the present topic is to 'Function effectively as an individual, and as a member or leader in diverse teams and in multi-disciplinary settings' (IEM, 2009). This is a challenge to education, because the dominant paradigm in tertiary education is individual learning and (especially) assessment. Student behaviour that in a university would be considered collusion, is in the profession considered collaboration. How then do educators develop the necessary team-based skills in graduates?

One way is through small individual design projects and other team-based exercises. However these are often only part of a course, and there is seldom reliance on team-projects for all the assessments in a course. Team conflict (Tuckman, 1965) is a difficult problem for educators to deal with, and they find it prudent to design assessments to avoid what looks like an unnecessary complication. But what if learning to deal with conflict and pressure was a *desirable* learning outcome? What would such a course look like? The closest are team-based project-courses, where the same team of student works on a project for the whole duration of the course.

In engineering this opportunity is provided by capstone final-year projects. All engineering degrees require such a course. It is usually a whole year in duration. In some universities it is undertaken individually, whereas in others it is done in teams. These team projects provide a setting in which students can develop their skills to work effectively as individual agents, while simultaneously having to coordinate their activities with other team members to accomplish larger outcomes.

Concept Model

Discrepancies in the Literature

As seen from the literature the relationship of the Five Factor Personality Model to academic performance is well known, however this relationship focusing on the education and profession of engineering is relatively unexplored.

With few exceptions (Rosander et al., 2011), there is a lack of quantitative assessments of the various factors that determine success in engineering education and how these relate to skills required in the profession. As explained in section 1.2 there is a notable difference between the skills required in the engineering profession and in engineering education. While an engineering degree consists mostly of engineering science-based papers and is assessed though standard practices such as exams and reports, the engineering profession as a practice consists primarily of project-based work involving multiple people, disciplines and expertise.

Purpose

There is a need to better understand how *personality* links to the development of *team skills* during engineering education. The purpose of this study is to explore the relationship between final year mechanical students' personality traits, their general academic performance in the final year of their degree and their performance in a capstone industry-based team-project.

Hypotheses

It was anticipated that academic performance would be affected by personality variables, general intelligence, and social skill which in turn influence job performance. This is represented as a hypothesised model in Figure 1.

From the study it was hoped that this model could be expanded upon and some of the relationships between variables be assessed and quantified.

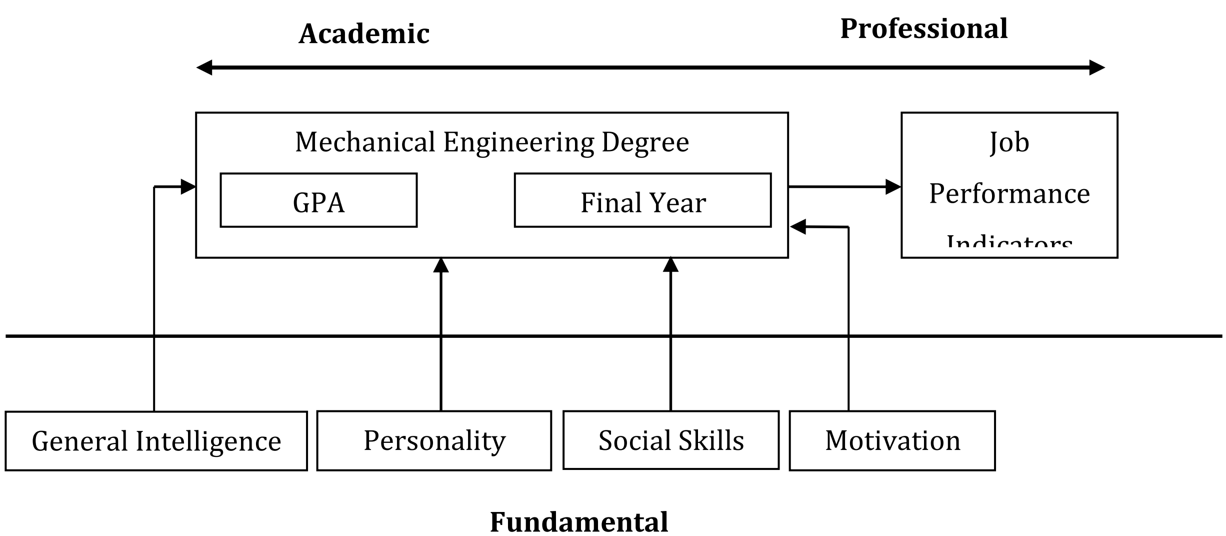


Figure 1: Hypothesised Concept model

Methodology

The overall approach was to survey engineering student-teams on a capstone course. There are many ways that universities run their final year projects, and not all have the same learning objectives in mind. This particular course is characterised by the following: full year course > students were placed into teams of predominantly four students with an academic staff member acting as supervisor > each team had a different project > projects were generally externally sponsored and therefore had a real industry client > allocation to teams was based on students' self-selection of project topic, not on self-section of team-members. This course design provides real engineering problems, real clients, and realistic team settings. It thus has a strong focus on developing the skills to transition from academic studies to professional practice. In contrast many universities offer one-student projects of internal origin, hence minimal team or professional interaction.

Procedure

An online survey was used in this study. It was sent to the participants at the beginning of the academic year and closed at its end. The majority of responses came soon after release or at closing. Those who responded to the survey filled out demographic questions as well as a personality test and gave consent for access to their academic records.

The participants were final year (4th year of study) mechanical engineering students at the University of Canterbury who were enrolled in the Final Year Projects course. The course consisted of 104 undergraduate mechanical engineering students, 88 male and 14 female. From these 43 usable participants were gained, 32 male and 11 female. The average age of usable participants was 22 years with a range of 20 to 45 (M = 22.00, SD = 3.79). Participation was voluntary with no material incentive given and the survey filled out in the students own time. Ethics approval was obtained from the University.

Measures

Personality test

The personality test used was 'A Very Brief Measure of the Big-Five Personality Domains' created by Gosling et al. (2003). This particular personality test measures the five super traits (Openness, Conscientiousness, Extraversion, Agreeableness and Emotional Stability), as opposed to the sub facets, through a 10 item measure. Although brief and therefore not as comprehensive as larger standard multi item instruments it is adequate as a measurement tool (Gosling et al., 2003). This test was chosen as it allowed for a timely response in the hopes that more participants would be obtained. This was considered important as the final year of engineering is renowned for a high work load and a lengthy survey generating negative word of mouth within the potential sample group was considered to be a significant detriment.

Represented as a model the five personality super traits which were found to have an effect on academic performance from this study are shown as a construct in Figure 2.

GPA

As a measure of general academic performance the participant's averaged grade point average (GPA) was used from the previous year and the first semester of the year that the survey was released. This data was extracted from the university records. The number of papers and the papers themselves that contribute to the measured GPA are the mostly the same for all the participants. Because of this GPA provides a reliable indicator of each participant's general academic performance comparable to the sample.

GPA is represented as a model with possible sub facets from the mechanical engineering curriculum, seen in Figure 3. Thus we are explicitly open to the possibility that different cognitive skills may be used in these different types of courses.

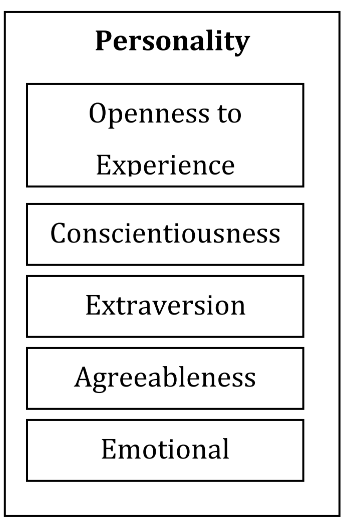


Figure 3: Modelled variable of Personality and super traits

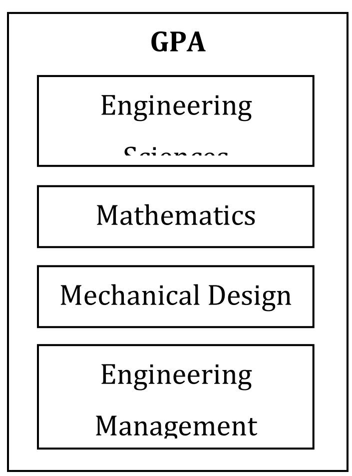


Figure 2: GPA consisting of various areas studied within mechanical engineering

Final Year Project

The full grade for the Final Year Project course consists of an individual mark and a team mark; both of these are moderated by the course director, and the team mark is determined by four assessors working independently. The Final Year Project course is double weighted (contributing twice as much to the honours grade than a normal paper) and therefore the students tend to put in significant effort into this paper compared to others.

The individual mark is primarily awarded by the teams' supervisor from a series of performance review meetings. It considers the workload undertaken by the particular student, their effectiveness as a team member, their ability to formally communicate, their application of knowledge and their ability to problem solve.

The team mark comes from the grading of written reports throughout the year, with the largest being the final report. These reports are marked by the course director, an independent marker, the supervisor, moderated by another academic, and commented on by the client. There is also an external moderation (sample basis). Assessment considers the overall quality of the reports, complexity of the problems undertaken, application of knowledge, problem analysis processes, means of investigation, and team effectiveness as a whole. Therefore team effectiveness is viewed by those awarding marks though the lens of various written reports.

Final year project marks including team and individual components are represented as a model in Figure 4, with the Team and Individual Mark contributing equally to the overall Final Year Project Grade.

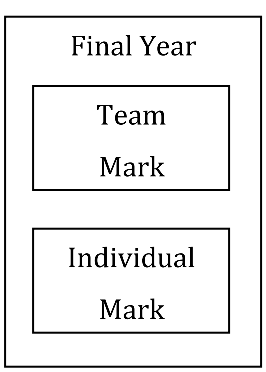


Figure 4: Marking for final year project (50% split)

Results

Statistical Results

Descriptive statistics for the sample group are presented in Table 3 along with a comparative sample of 1126 Caucasian undergraduate psychology students from the University of Texas at Austin using the same personality test (Gosling et al., 2003).

The personality results for this group of engineering students fit with what has previously been identified as typical traits for engineers as described in section 1.2, that is, high conscientiousness and emotional stability, reasonable agreeableness and extraversion.

Agreeableness and Emotional Stability have statistically significant differences between the two samples, $t=3.706$ at $p< 0.01$, and $t = 2.695$ at $p<0.01$ respectively. The difference between personalities is similar to traits typical of engineers as described by Van Der Molen et al. (2007), specifically *lower Agreeableness and higher Emotional Stability compared to the general population*. However this is simply a comparison between mechanical engineering students in New Zealand and psychology students in Texas, therefore conclusions regarding a typical personality type for engineers should not be made.

Table 3: Descriptive statistics of variables

Variables	Study Results (N=43)		Comparative results (N=1126)	
	Mean	**Std. deviation**	**Mean**	**Std. deviation**
Openness	5.407	0.915	5.43	1.06
Conscientiousness	5.779	1.135	5.47	1.13
Extraversion	4.849	1.303	4.56	1.48
Agreeableness	4.651	1.055	5.26	1.12
Emotional Stability	5.326	1.123	4.85	1.45
GPA	6.413	1.516		
Individual Mark	79.179	10.258		
Team Mark	79.302	10.296		

Note: the competitive results are from (Gosling et al., 2003). Significance results shown in grey, $p>0.01$

The inter-correlations with significance for this study are given in Table 4. Compared to those from meta-analyses of personality to academic performance, seen in Table 2, Conscientiousness in both sets of results was the strongest predictor of performance with a positive correlation. Both sets of results also had weak correlation between the other super traits and academic performance. Despite the statistical insignificance it is noted that the correlations for Openness and Extraversion are opposite to previous results seen in Table 2.

Table 4: Correlations and significance of variables, correlations greater than ±0.1 and significant results are bolded

	1	2	3	4	5	6	7	8
	O	C	E	A	ES(N)			
1. Openness	-							
2. Conscientiousness	**-0.135** (0.388)	-						
3. Extraversion	-0.042 (0.789)	-0.011 (0.944)	-					
4. Agreeableness	-0.078 (0.621)	0.053 (0.734)	-0.083 (0.599)	-				
5. Emotional Stability	-0.078 (0.621)	0.016 (0.920)	**0.218** (0.161)	**0.234** (0.131)	-			
6. GPA	-0.057 (0.718)	**0.307 (0.045)**	**0.176** (0.260)	0.049 (0.757)	0.092 (0.557)	-		
7. Individual Mark	**0.152** (0.329)	0.057 (0.715)	0.096 (0.542)	**0.148** (0.343)	**-0.131** (0.404)	**0.396 (0.008)**	-	
8. Team Mark	0.096 (0.542)	-0.061 (0.698)	0.010 (0.951)	-0.028 (0.856)	**-0.195** (0.211)	**0.435 (0.004)**	**0.829 (0.000)**	-

Note: significance given in brackets, $p < 0.05$.

Further regression analyses did not reveal any deeper relationship than those shown in Table 4.

Model

From the survey results aspects of the hypothesis model seen in Figure 2, were able to be expanded on and some relationships quantified. Figure 5 shows the results model which includes the variables assessed in the survey that had correlations of greater than 0.10. The percentages shown on the model are the variances given as a percentage (r^2 X 100), this percentage is the amount of variation one variable accounts for in another.

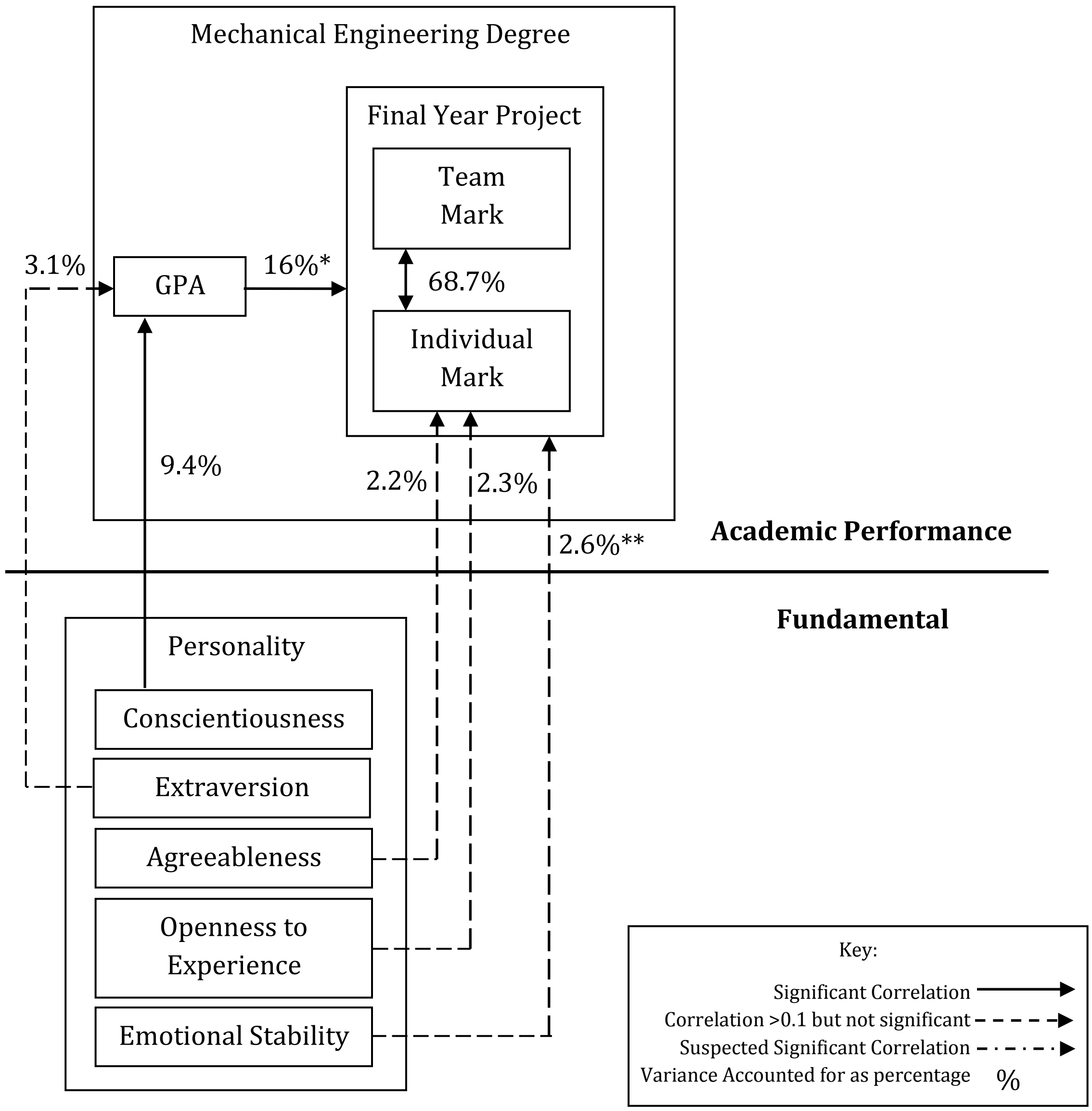


Figure 5: Results Model

* The correlative values GPA has with Team Mark and Individual Mark have been averaged together

** The correlative values Emotional Stability has with Team Mark and Individual Mark have been averaged together

A student's *Conscientiousness was found to account for 9.4% variance of their GPA* assuming that it was the driver of GPA, while extraversion accounted for 3.1% of variance within GPA. GPA could account for 16% of the variance for the total project mark assuming that was the driver of project mark. Within the project team mark and individual mark combined accounted for 68.7% of variance between each other; both team and individual marks were earned simultaneously so neither was an obvious driver.

Discussion

Outcomes

What are the drivers for success in this type of capstone engineering learning environment?

As was hypothesised and predicted by previous studies *Conscientiousness had a significant correlation with GPA*, although unexpectedly *not with marks for the Final Year Project*. This may be explained through changes in motivation. It has been found that Conscientiousness is positively correlated with various forms of academic motivation (De Feyter et al., 2012; Hazrati-Viari et al., 2012; Komarraju & Karau, 2005; Komarraju, Karau, & Schmeck, 2009). A student may be able to find motivation based on their own conscientiousness for GPA related academic performance. In the Final Year Project *external factors may have a stronger influence on academic motivation*, namely the supportiveness and/or demands of the team, supervisor and client.

The weaker influences of Agreeableness and Emotional Stability are of interest although they are not statistically significant. It is intuitive that Openness and Agreeableness influence the individual mark with positive correlations as a fair amount of this mark is based on the supervisors' assessment of the student through interviews. A student with high Openness may come across as being more intelligent and a high level of Agreeableness may make conversation with them more pleasant and constructive, resulting in a better Individual Mark.

The negative correlation of Emotional Stability with both project marks is opposite to most previous studies (although most studies also have weak correlations) in which students who have greater Emotional Stability tend to have better academic performance (O'Connor & Paunonen, 2007; Poropat, 2009). This study suggests that *students who experience some anxiety regarding performance in the Final Year Project will perform better*, a concept which has been suggested previously (De Feyter et al., 2012; Komarraju & Karau, 2005; Komarraju et al., 2009).

Engineers require not only intellect, knowledge and the ability to apply these to problem solving, but also interpersonal skills. The results indicate that there are significant other variables which account for variability in academic performance; it is likely that interpersonal and social skills are included in these unknowns as identified in Figure 6. It is suspected that the ability for a student to work with their peers in all manner of measured academic tasks such as lab, test and exam preparation and report writing will give a significant advantage.

Comprehensive Model

Based on previous literature and the results of this study the hypotheses concept model, seen in Figure 1, has been greatly expanded upon. This comprehensive model is seen in Figure 6.

In previous studies general intelligence as well as fluid, crystallised, and other facets of intellect have linked to academic performance (A. Furnham et al., 2005; Moutafi et al., 2003, 2006, 2004; Moutafi, Furnham, & Paltiel, 2005). Motivation as a driver for performance was not explored though the survey but is likely to be a powerful predictor variable. It is likely that a person's social skills would have an influence on their academic performance, particularly in the final year project and other team based course work; Positive relationship between social adjustment and intelligence have been found (Austin et al., 2002). There are likely to be many other factors also influencing academic performance which have not been considered at this point.

Academic performance was measured holistically but could be divided into sub facets based on areas within the mechanical engineering curriculum. These sub facets would tend to have different performance measurement criteria (courses may be exam and test heavy as oppose to report or presentation heavy) and particular skills and knowledge from each may relate differently to other variables.

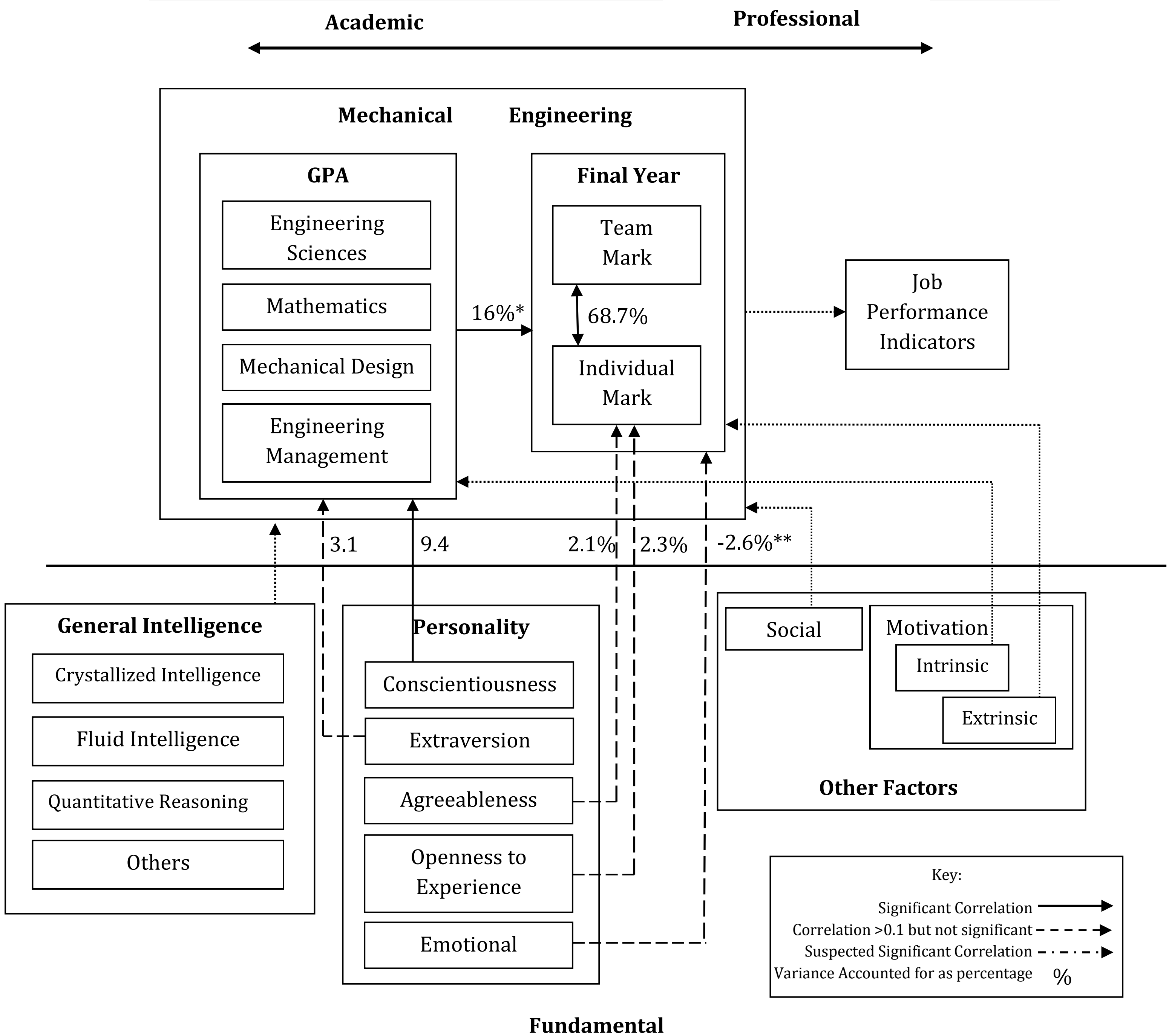


Figure 6: Comprehensive Model

* The correlative values GPA has with Team Mark and Individual Mark have been averaged together

** The correlative values Emotional Stability has with Team Mark and Individual Mark have been averaged together

Implications for Practitioners

The results of this study can affect several practitioner groups in engineering education and possibly further afield. Key groups indentified are the students, the educators and the final group is team managers in industry.

Students that undertake a final year project as described in this paper or a similar research and development based team project are primarily interested in passing the course, so their marks and grades are important to them. There are three implications identified for such students.

The personality trait Conscientiousness is associated with overall GPA. This implies that the ability to will oneself to have a strong work ethic is beneficial for Academic Performance, a well-known phenomenon. Interestingly, GPA and Extraversion were positively correlated with each other, although not significantly, this may be explained by the need for students to be outgoing enough to approach each other and their lectures for help.

Conscientiousness is not associated with the Project Mark but is with GPA. This implies that there are other factors driving work ethic with regards to the project (likely external demands from the team, supervisor and client) and that knowledge and skills learnt in other courses are beneficial to the total Project Mark.

Individual Mark on the project course is weakly associated with higher Openness, higher Agreeableness, and a lower Emotional Stability. This may be explained as the Individual Marks are assessed though formal interviews with supervisors as well as informal contact throughout the year with the supervisor and client. A person who is open to new ideas, agreeable and slightly neurotic about the project suggesting that they are passionate about it may be more favourably looked upon by markers.

Team Mark in the Final Year Project is associated with lower Emotional Stability of the individuals within the team. Lower Emotional Stability relating to Team Mark is an interesting and unexpected association. This can be interpreted that a small degree of anxiety regarding the Project will increase student's intrinsic academic motivation. This study has not assessed motivation and the extent that emotional stability, or any other personality trait, may influence it.

The second group is engineering practitioners. As discussed above Final Year Projects that generate a small degree of anxiety in students may have a better academic performance. The

Final Year Projects in this particular study generally had an external (to the university) industry client; it is possible that the presence of this stakeholder may generate the required anxiousness to succeed. Few other universities offer an engineering degree with a capstone project that offers such a link with industry.

The third group of practitioners are managers of design teams in industry. High GPA is associated with individual and team performance. This implies that although GPA may be helpful as an indicator of job performance it is important to note that GPA only accounts for approximately 16% of the variance in Individual and Team Marks. Other variables affecting performance in the Final Year Project, which is intended to represent industry work, have not been determined in this study.

Limitations

This study had several limitations, the following have been identified as areas that may improved upon in future studies and/or are likely to have had a significant impact on the results.

The sample size used in this study was relatively small compared to similar studies examining personality and academic performance. A larger sample size may have provided more conclusive results but despite this strong correlations were still found. Although this sample size was small compared to the general population it was near half the enrolled students in this particular year group studying mechanical engineering at the University of Canterbury. Therefore this information may prove particularly useful for stakeholders requiring insight into these specific students.

The Personality test used was an adequate tool to measure the five super traits (Gosling et al., 2003). A more comprehensive personality test may have provided stronger evidence for correlations between personality and academic performance as well as giving insight into possible links between the sub facets of the five factor model and academic performance.

It was hoped in this study that there were be a reasonable number of teams where all members had completed the survey, unfortunately only a small number had. Complete teams would allow a comparison of personalities within and between teams. This may give insight into conflict within teams and ideal personality combinations within teams affecting academic performance with regard to the Final Year Project.

The survey used to gather data was opened at the start of the academic year and closed at its end, with students participating in the survey predominately at opening and closing. Because of this personality results may not be consistent between the students as it is possible that the high work load of their final year may have adjusted their answers to the personality questionnaire similar to changes in personality over time in medical students (Lievens, Coetsier, De Fruyt, & De Maeseneer, 2002). In addition the students were able to fill out the survey in their own time, meaning that it may have been done in a variety of scenarios, affecting answers. Students who answered the personality test at the university may have given different answers than they would have out of university setting.

Final Year Projects varied significantly between teams in terms of supervisor's clients and the knowledge and skills used in the project itself. Because of this there is an inherent difficulty comparing project marks but as the grades that were used in this study are the same as those that are used by the university to award honours grades they are assumed to adjust for reasonable comparability.

The GPA marks used to represent academic performance were earned by the students over a number of different courses. There were differences in the courses done between students, but within the engineering degree the majority of papers are common between all students.

Direction of Future Research

From the comparison between the results model, seen in Figure 5, and the comprehensive model, seen in Figure 6 there are many other variables that may account for academic performance.

From this study it was indicated that there were potentially significant correlations between Academic Performance, and Agreeableness and Emotional Stability. Specifically a lower Emotional Stability associated with a better individual and team performance, as well as Extraversion and GPA.

A future survey might consider further investigating the link between the sub facets of the personality super traits, motivation and academic performance. Taking into account the satisfaction of students within their teams, this may give insight into the social factors affecting academic performance.

Conclusions

This paper makes several novel contributions. First, it specifically examines the role of personality in team-projects for engineering capstone courses. These are important courses in the educational programme for engineers, and there are high expectations from the external profession regarding the outcomes of such courses, so it is worth better understanding the educational dynamics. The work confirms some known associations between personality and academic success, disconfirms others, and in yet other cases finds tentative new associations.

Second, this paper provides a structured model of proposed causality, that links personality, teams, and performance on a capstone industry-based team-project. This is important, because it suggests implications whereby the efficacy of the professional graduate attribute may be enhanced.

To summarise the main statistical findings: Openness and Agreeableness influence the individual mark, and students who experience some mild anxiety (Neuroticism) regarding performance in the Final Year Project perform better. Conscientiousness affects overall GPA but not project mark.

Contribution statement

The research was primarily conducted by RH, who investigated the literature, designed the survey, obtained the ethics approval, and analysed the results. DP set the research agenda, supervised and guided the research, provided the professional engineering context, and provided access to student participants. WH provided guidance and specialist psychology knowledge, and contributed to the data analysis. All authors contributed to the writing of the paper.

References

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*(2), 219–245. doi:10.1037/0033-2909.121.2.219

Allik, J., & Realo, A. (1997). Intelligence, academic abilities, and personality. *Personality and Individual Differences*, *23*(5), 809–814. doi:10.1016/S0191-8869(97)00103-7

Austin, E. J., Deary, I. J., Whiteman, M. C., Fowkes, F. G. ., Pedersen, N. L., Rabbitt, P., Bent, N., et al. (2002). Relationships between ability and personality:does intelligence contribute positively to personal and social adjustment? *Personality and Individual Differences*, *32*(8), 1391–1411. doi:10.1016/S0191-8869(01)00129-5

Balsamo, M., Lauriola, M., & Saggino, A. (2012). Personality and college major choice: which come first? *Psychology (Irvine)*, *3*(5), 399+.

Banik, G. (2008). *Employer's expectations of the performance of CM graduates.* Paper presented at the ASEE Conference, Chantilly, VA 20153, United States.

Barrick, M. R., & Mount, M. K. (1991). THE BIG FIVE PERSONALITY DIMENSIONS AND JOB PERFORMANCE: A META-ANALYSIS. *Personnel Psychology*, *44*(1), 1–26. doi:10.1111/j.1744-6570.1991.tb00688.x

Brand, C. R. (1994). Open to experience—closed to intelligence: Why the “Big Five” are really the “Comprehensive Six.” *European Journal of Personality*, *8*(4), 299–310. doi:10.1002/per.2410080407

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, *44*(7), 1596–1603. doi:10.1016/j.paid.2008.01.003

Chamorro-Premuzic, T., Moutafi, J., & Furnham, A. (2005). The relationship between personality traits, subjectively-assessed and fluid intelligence. *Personality and Individual Differences*, *38*(7), 1517–1528. doi:10.1016/j.paid.2004.09.018

De Feyter, T., Caers, R., Vigna, C., & Berings, D. (2012). Unraveling the impact of the Big Five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learning and Individual Differences*, *22*(4), 439–448. doi:10.1016/j.lindif.2012.03.013

Demetriou, A., Kyriakides, L., & Avraamidou, C. (2003). The missing link in the relations between intelligence and personality. *Journal of Research in Personality*, *37*(6), 547–581. doi:10.1016/S0092-6566(03)00048-5

Dickson, B., & Grant, C. D. (2008). *Engineering management skills: The missing link for technical graduates.* Paper presented at the ASEE Conference, Chantilly, VA 20153, United States.

Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, *41*(1), 417–440. doi:10.1146/annurev.ps.41.020190.002221

Ditcher, A. K. (2001). Effective Teaching and Learning in Higher Education, with Particular Reference
to the Undergraduate Education of Professiona Engineers. *International Journal of Engineering Education*, *17*(1), 24-29.

Furnham, A., Forde, L., & Cotter, T. (1998). Personality and intelligence. *Personality and Individual Differences*, *24*(2), 187–192. doi:10.1016/S0191-8869(97)00169-4

Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and Intelligence: Gender, the Big Five, Self-Estimated and Psychometric Intelligence. *International Journal of Selection and Assessment*, *13*(1), 11–24. doi:10.1111/j.0965-075X.2005.00296.x

Furnham, A.-P. (2004). Estimating one’s own personality and intelligence scores. *British Journal of Psychology*, *95*(2), 149–160.

Futrell, G. (1985). What Do You Look for in a New Engineer? *Journal of Management in Engineering*, *1*(1), 20–27. doi:10.1061/(ASCE)9742-597X(1985)1:1(20)

Gosling, S. D., Rentfrow, P. J., & Swann Jr., W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. doi:10.1016/S0092-6566(03)00046-1

Harrison, R., Tomblen, D. T., & Jackson, T. A. (1955). Profile of the Mechanical Engineer III. Personality. *Personnel Psychology*, *8*(4), 469–490. doi:10.1111/j.1744-6570.1955.tb01224.x

Hazrati-Viari, A., Rad, A. T., & Torabi, S. S. (2012). The effect of personality traits on academic performance: The mediating role of academic motivation. *Procedia - Social and Behavioral Sciences*, *32*(0), 367–371. doi:10.1016/j.sbspro.2012.01.055

IEM (2009). Graduate Attributes and Professional Competencies Version 2 - 18 June 2009. Retrieved 23 Nov 2009, from http://www.washingtonaccord.org/IEA-Grad-Attr-Prof-Competencies-v2.pdf

IPENZ (2006). MANUAL FOR THE ACCREDITATION OF PROFESSIONAL ENGINEERING AND ENGINEERING TECHNOLOGY PROGRAMMES 5. Retrieved 12 June 2009, from http://www.ipenz.org.nz/IPENZ/Forms/pdfs/Accreditation-Manual-5th-Edition-November-06.pdf

Kappe, R., & van der Flier, H. (2010). Using multiple and specific criteria to assess the predictive validity of the Big Five personality factors on academic performance. *Journal of Research in Personality*, *44*(1), 142–145. doi:10.1016/j.jrp.2009.11.002

Kichuk, S. L., & Wiesner, W. H. (1997). The big five personality factors and team performance: implications for selecting successful product design teams. *Journal of Engineering and Technology Management*, *14*(3–4), 195–221. doi:10.1016/S0923-4748(97)00010-6

Kline, P., & Lapham, S. L. (1992). Personality and faculty in British universities. *Personality and Individual Differences*, *13*(7), 855–857. doi:10.1016/0191-8869(92)90061-S

Komarraju, M., & Karau, S. J. (2005). The relationship between the big five personality traits and academic motivation. *Personality and Individual Differences*, *39*(3), 557–567. doi:10.1016/j.paid.2005.02.013

Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, *19*(1), 47–52. doi:10.1016/j.lindif.2008.07.001

Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: a five-factor model perspective. *Medical Education*, *36*(11), 1050–1056. doi:10.1046/j.1365-2923.2002.01328.x

McCrae, R. B., & Costa, P. T. (1990). *Personality in adulthood*. New York: Guilford.

Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: a study using the Neo Personality Inventory and the Myers–Briggs Type Indicator. *European Journal of Personality*, *17*(1), 79–94. doi:10.1002/per.471

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences*, *16*(1), 31–42. doi:10.1016/j.lindif.2005.06.003

Moutafi, J., Furnham, A., & Paltiel, L. (2004). Why is Conscientiousness negatively correlated with intelligence? *Personality and Individual Differences*, *37*(5), 1013–1022. doi:10.1016/j.paid.2003.11.010

Moutafi, J., Furnham, A., & Paltiel, L. (2005). Can personality factors predict intelligence? *Personality and Individual Differences*, *38*(5), 1021–1033. doi:10.1016/j.paid.2004.06.023

NEEP Final Project Report, 2010. *National Engineering Education Plan*. http://www.ipenz.org.nz/ipenz/education_career/NEEP_Project_Report.pdf

Newport, C. L., & Elms, D. G. (1997). Effective Engineers. *International Journal of Engineering Education*, *13*(5), 325-332.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, *43*(5), 971–990. doi:10.1016/j.paid.2007.03.017

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. doi:10.1037/a0014996

Rosander, P., Bäckström, M., & Stenberg, G. (2011). Personality traits and general intelligence as predictors of academic performance: A structural equation modelling approach. *Learning and Individual Differences*, *21*(5), 590–596. doi:10.1016/j.lindif.2011.04.004

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63*(6), 384-399.

Van Der Molen, H. T., Schmidt, H. G., & Kruisman, G. (2007). Personality characteristics of engineers. *European Journal of Engineering Education*, *32*(5), 495–501. doi:10.1080/03043790701433111

Wadhwa, L. C. (1981). *Management practices survey in Australian companies.* Paper presented at the Conference on Engineering Management, Barton, ACT, Australia.

A 2 Usability Analysis Methods Review

In addition to measuring the participant's severity of impairment using the ICF, a means to measure usability of the controller system was needed to make accurate comparisons.

In this section methods to measure aspects of usability which were explored but ultimately discarded are overviewed. This is followed by the chosen methods of the System Usability Score (SUS), NASA Task Load Index (TLX) and a simple close out interview.

To best asses usability in the study the number of participants were considered a key variable. It is touted that as low as five appropriate users are needed to determine usability (Lewis, 1994; Jakob Nielsen, 1994; Virzi, 1992). For purposes of this study a higher number of participants was needed to be able to conduct strong statistically significant assessments.

A 2.1.0 Considered Measurement Methods

Before settling on quantifying only the users impairment (or lack of) using ICF, several other methods were explored that could be used instead of, or in addition to ICF. These methods primarily focused on quantifying the actions that the users needed to make (Fitts and Hicks Law) or quantified the controller (Complexity, Usability and Information Entropy).

These methods involved utilising measurements of ability such as movement speed, cognitive ability in relation to complexity of task, and visibility and clarity of visual information as relating to established measurements of interface. Specific methods examined were as follow:

Hierarchical Task Analysis and Critical Path Analysis

Hierarchical Task Analysis and its sub analysis technique Critical Path Analysis applied in a usability context describes a task by breaking it into a hierarchical structure of goals, sub-goals, operations and plans. Combining this task break down into single actions with a rough quickest possible time taken to complete each action, would allow an optimum time to be determined (Crandall, Klein, & Hoffman, 2006; B. Kirwan & Ainsworth, 1992; Lewis, 1994; Polson, Lewis, Rieman, & Wharton, 1992; N. A. Stanton, 2006; P. N. A. Stanton, Walker, & Salmon, 2013; Stewart, 1992).

Recording the time taken for each action done by each participant would have allowed comparison to this optimum. However tracking and recording participant's exact actions proved impractical as further explained in section 3.3.1 Methodology Observation Methods.

Fitts's Law and Hick's Law

Fitts's Law predicts that the time required to rapidly move to a target area is a function of the distance to the target and the size of the target (Fitts & Peterson, 1964). In this study it is the time taken to reach for and press a button. However Fitt's law is primarily a test of reaction time and does not consider a user's cognitive processing time and a lack of urgency, these criteria are particularly problematic when considering people with impairments (Chapuis, Blanch, & Beaudouin-Lafon, 2007; Keele, 1968; Vella, Vigouroux, & Gorce, 2009; Wobbrock, Cutrell, Harada, & MacKenzie, 2008).

Hicks's Law describes the time taken for a person to make a decision as a result of the possible choices they have: increasing the number of choices will increase the decision time logarithmically (Hick, 1952). For practical and accurate application of Hick's Law users are required to have predictable levels of performance, be involved in a speed accuracy trade off task and most importantly have predictable levels of intelligence (Beh, Roberts, & Prichard-Levy, 1994; Lindley, Bathurst, Smith, & Wilson, 1993; Roberts, Beh, & Stankov, 1988; D. W. Schneider & Anderson, 2011; Usher, Olami, & McClelland, 2002). The industry focused aspect of this project requires a means to easily determine the abilities of a user, measuring intelligence and assuming consistent circumstances could only be practically done in the lab and would only aid a small aspect of the study.

These laws would have been best applied in optimising a controller design as opposed to gaining a general understanding of user related measures affecting performance. Work by others has been done to combine the laws for practical application (Beggs, Graham, Monk, Shaw, & Howarth, 1972), however this work has not reached the point for simple use in a study such as this.

Complexity and Information Entropy Measures

Several methods were explored and tested to quantify how easy the controller was to use. A method of measuring complexity based on the Shannon formula used in communication theory was a potential means to predict the usability of a screen (Comber & Maltby, 1996, 1997, 1995). This method was applied to the controller menu structure where complexity was determined based on the number of icons present, their relative sizes and arrangement to each other (Comber & Maltby, 1996, 1997, 1995). This proved inaccurate in the context of the study and impractical for industry application as the menu structure and icon arrangement were not applicable for the method which was still in development.

Direct measures of screen aesthetics and usability guidelines were applied as a means to determine the controllers global usability (Ngo, Teo, & Byrne, 2000; J. Nielsen, 1993; Jakob Nielsen, 1994; D. A. Norman, 1983a, 1983b; Seffah, Donyaee, Kline, & Padda, 2006). However these methods could not be easily integrated with quantification of the user's abilities to give realistic predictions of performance.

The principle of entropy in information theory (Shannon entropy) was also applied to determine the expected value of information presented to participants, with the expectation that it would correlate with performance and therefore act as a usability measure (Bonsiepe, 1968; Comber, 2010; Maltby, 2007; Shannon, 2001; Shannon & Weaver, 1962). Application of information theory proved possible but many variables were subjective and the use of it beyond this study would not have been practical.

A 2.2.0 Considered Survey Methods

The most practical and effective means to determine usability was to give participants simple standardised usability questionnaires that had proven use in practice.

Standardised usability questionnaires looked at but not used were:

Given before the experiment

- The Computer Attitude Measure (CAM) (Kay, 1993a)
- The Computer Ability Survey (CAS) (Kay, 1993b)

Given directly after a task

- After-scenario Questionnaire or ASQ (Lewis, 1991, 1995)
- Expectation Ratings or ER (Albert & Dixon, 2003)
- Usability Magnitude Estimation or UME (McGee, 2003, 2004)
- Single Ease Question or SEQ (Tedesco & Tullis, 2006)
- Subjective Mental Effort Question or SMEQ (Sauro & Dumas, 2009)

Given at the end of the experiment

- Questionnaire for User Interaction Satisfaction or QUIS(Chin, Diehl, & Norman, 1988)
- Software Usability Measurement Inventory or SUMI (Kirakowski & Corbett, 1993; McSweeney, 1992)
- Post-study System Usability Questionnaire or PSSUQ (Lewis, 1992, 1995, 2002)
- Subjective Workload Assesment Technique or SWAT (Reid & Nygren, 1988)
- Workload Profile or WP (TSANG & VELAZQUEZ, 1996)

The chosen end of the experiment and end of task questionnaires were the System Usability Scale (SUS) and Task Load Index (TLX) respectively.

The selected SUS is short, simply worded and is broadly applicable, the TlX was chosen as it captures a more emotive response (dislike, frustration etc) and is also short and simply worded. The SUS and TLX are reviewed further as follows.

A 2.2.1 System Usability Scale

For this study a means to measure the user's opinion of the controller was needed that would quantify these opinions, was simple to implement and usable beyond this study should industry want to apply it. The System Usability Scale (SUS) was identified as meeting these criteria.

The SUS was developed in the mid 1980's by John Brooke and released in 1986 informally by Brooke, it has since reached status of a de facto standard and has been in more than 1,200 publications from its release to 2013 (Brooke, 2013).

It was intended to be a 'quick and dirty' means to determine usability from the user's perspective (Brooke, 1996). Over the years SUS has proven itself not be 'dirty' but a reliable method for usability analysis in a variety of situation (Bangor, Kortum, & Miller, 2008; Borsci, Federici, & Lauriola, 2009; Lewis, 1995; Lewis & Sauro, 2009; Sauro & Lewis, 2009).

The SUS uses a ten-item attitude Likert scale (questions answered Strongly Agree to Strongly Disagree) to measure usability in terms of effectiveness, efficiency, and satisfaction. The SUS is scored out of 5 (with 5 as maximum usability), an average score is considered to be 3.4 based on 5000 SUS observations (Sauro, 2011). Its application is explained in section 3.3.2 Methodology – Participant Assessments.

A 2.2.2 Task Load Index

The Task Load index (TLX) was developed by NASA (hence it is often called the NASA-TLX), it was released for publication in 1988 and is a subjective, multidimensional assessment tool that rates self reported perceived workload and performance of a task by a user (Hart & Staveland, 1988).

The TLX consist of six questions that are answered on a 20 point scale. These questions measure Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration.

Since its release the TLX has been used extensively in a range of applications, many of which are similar to this study and since it's release it has been reviewed independently and updated for modern use (Hart, 2006; Noyes & Bruneau, 2007; Rubio, Díaz, Martín, & Puente, 2004).

A 3 Cognitive Task Analysis Word List

Ankle
Saloon
Icebox
Slipper
Infant
Mucus
Pudding
Hostage
Banner
Bullet
Sulphur
Doorman
Locker
Piano
Sunburn
Missile
Thicket
Monarch
Cowhide
Leopard
Bath
Castanet
Chef
Crayon
Goose
Graphic
Hovercraft
Operation
Dill
Fairies
Level
Shears
Accountant
Bag
Brochure
Distance
Locket
Pair
Range
Dashboard
Hub
Purchase
Rectangle

Piston
Butcher
Fiord
Typhoon
Nectar
Harness
Reptile
Lobster
Rattle
Bandit
Pepper
Morgue
Trumpet
Singer
Blister
Jelly
Salad
Settler
Sultan
Fabric
Back
Play
Giants
Size
Beginner
Cherry
Fowl
Account
Mountain
Wave
Produce
Powder
Carpenter
Competition
Voyage
Scale
Curtain
Worm
Bee
Approval
House
Bulb
Grain

Lemon
Hamlet
Shotgun
Abode
Poster
Cigar
Painter
Steamer
Sunset
Costume
Bagpipe
Banker
Spinach
Hairpin
Beggar
Skillet
Invoice
Robber
Kettle
Glacier
Cannon
Country
Horses
Line
Care
Pull
Smile
Ants
Bottle
Sky
Baseball
Notebook
Rose
Pollution
Book
Icicle
Rifle
Star
Way
Trip
Cave
Eggs
Sink

A 4 University of Canterbury Ethics Application

UNIVERSITY OF CANTERBURY

HUMAN ETHICS COMMITTEE

APPLICATION FOR REVIEW & APPROVAL

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

This form should be completed in the light of the Principles and Guidelines issued by the Human Ethics Committee. Applicants must read those before filling out the application form. The latest versions of both the Guidelines and the Application Form can be found on the website of the Human Ethics Committee.
website: http://www.canterbury.ac.nz/humanethics

This application form is to be used for Applications *NOT* covered by the Educational Research Human Ethics Committee (ERHEC)

NOTE: This electronic copy may not have sufficient space for completion of all parts of the form if downloaded as a blank copy of the application form. It is intended as a template for use by those staff and students who have access to a word processor. When typing in please type where the paragraph marks start after each question, not in the actual boxes.

Please submit **SIX printed or typed copies** and **ONE electronic copy** of the completed application duly signed by applicant and supervisor or Head of Department, and all relevant documents referred to in questions 3, 7, 8, 9, 10, 11, 15 (i.e. authorizations, approvals, information and consent forms). Hard copies should be sent to the Secretary, Human Ethics Committee, Okeover House and electronic copies to human-ethics@canterbury.ac.nz. Please note that it is preferred the electronic copy to be one document signed, scanned and forwarded to the Secretary of the Human Ethics Committee.

1. PROJECT NAME: Methodology for identifying Errors and Efficiency in Users Interfaces with Regard to User Ability: Error Tracking on Wheelchair Interfaces

2. NAME OF APPLICANT: Rory Horne

Contact Telephone No: 0273521678

UNIVERSITY DEPARTMENT (or other contact address): Mechanical Engineering

EMAIL ADDRESS: rory.horne@pg.canterbury.ac.nz

STATUS OF PROJECT: ME (Masters of Engineering)

NAME OF SUPERVISOR: Dirk Pons, Deak Helton

OTHER INVESTIGATORS: Dynamic Controls

SIGNED BY: Applicant: .... Date: 1/07/2013

HOD/Supervisor: .... Date: 28/6/13

The checklist on the following page must be completed and signed by the applicant and, if the applicant is a student, by the applicant's supervisor

CHECK LIST

Please check the following items before sending the completed form to the Committee.

All the necessary signatures on page 1 have been obtained.	**[ ]**
All the necessary approvals under Question 3 have been obtained or are the subject of correspondence of which copies are attached.	**[ ] or NA**
A copy of any questionnaire, with an appropriate rubric at the beginning or accompanied by an appropriate covering page, is attached.	**[ ] or NA**
A list of interview topics and, for a structured interview, a detailed list of questions, is attached.	**[ ] or NA**
A copy of any advertisement, or notice, or informative letter asking for volunteers is attached.	**[ ] or NA**
A copy of each information sheet required is attached.	**[ ] or NA**
A copy of each consent form required is attached.	**[ ] or NA**
A copy of the required debriefing sheet is attached.	**[ ] or NA**
An electronic copy of the signed application has been forwarded to the HEC	**[ ] or NA**

Attention to the preceding check list is intended to ensure that the application and its documentation have been thoroughly reviewed by the applicant and (where applicable) by the supervisor and that the preparation of the project is up to the standard expected of and by the University of Canterbury.

The signature of the applicant will be understood to imply that the applicant has designed the project and prepared the application with due regard to the Principles and Guidelines of the HEC, that all the questions in the application form have been duly answered and that the necessary documentation has been properly formulated and checked.

Signature of Applicant 1/07/2013

The signature of the supervisor will be understood to imply in addition that, in the judgment of the supervisor, the design and documentation are of a standard appropriate for a research project carried out in the name of the University of Canterbury or for training in such research.

Signature of Supervisor

26 June 2013

		Delete whichever is in-applicable
3 (a)	WILL THE PROJECT REQUIRE ETHICAL APPROVAL FROM OTHER BODIES? eg Health and Disability Ethics Committee (HDEC) If Yes, please explain how this approval has been or will be obtained, enclosing copies of relevant correspondence. *NOTE: To save time, it is recommended that in the case of HDEC applications, an application is made concurrently with the application to the UC HEC.*	**Yes**

An application was begun to the HDEC but the following message was recived *"According to your answers above, your study does not require HDEC review"* on page 9 of 88 on the online application, in the 'inclusions' section.

The HDEC application form is included as a separate document in this application.

(b)	WILL THE PROJECT REQUIRE APPROVAL FOR ACCESS TO THE PARTICIPANTS FROM OTHER INDIVIDUALS OR BODIES? (e.g., parents, guardians, school principals, teachers, boards, responsible authorities including employers, etc.) If Yes, please explain how this approval has been or will be obtained, enclosing copies of relevant correspondence.	**No**
(c)	WILL THE PROJECT REQUIRE MAORI CONSULTATION? If Yes, please provide evidence that consultation has occurred or, if underway, provide a copy of approval once gained.	**No**
(d)	WILL THE PROJECT REQUIRE COMMUNITY CONSULTATION? If Yes, please provide evidence of appropriate consultation.	**No**
4 (a)	IS THE PROJECT BEING EXTERNALLY FUNDED? If Yes, please identify the source of funds.	**Yes: MSI (Ministry of Science and Innovation)**
(b)	IS THE PROJECT COMMISSIONED BY OR CARRIED OUT ON BEHALF OF AN EXTERNAL BODY? If Yes, please identify the body and any Intellectual Property agreements. This includes ownership of data and reports arising.	**Yes**

Dynamic Controls is the industry partner associated with this project. It has been agreed that Dynamic Controls will receive a copy of the thesis containing the processed results of the project. They will not receive the raw data, video recordings collected or any personal information of the participants.

(c)	IS THE PROJECT TO BE PART OF THE CEISMIC DIGITAL ARCHIVE? If so, please ensure all participants are made aware of this, and have filled in the UC CEISMIC Quake Studies consent form. See www.ceismic.org.nz.	**No**

Further, please ensure that all participants are made aware of any of the above in information sheets and consent forms provided.

include any personal participant details.

or (b) The research is not anonymous, but is confidential and informed consent will be obtained through a signed consent form (include a copy of the consent form and information sheet) **Yes**
This means that while you do/may know the identity of the participants, with respect to the data provided, you will not make their identity public.

Where confidentiality is promised, what will be done to ensure that the identities of participants cannot be known by unauthorized persons? (e.g. use of pseudonyms and disguising of identifying material).

or (c) The research is neither anonymous nor confidential and informed consent will be obtained through a signed consent form (include a copy of the consent form and information sheet). **No**

or (d) Informed consent will be obtained by some other method – please specify and provide details. **No**

A copy of the consent form for participants is included as a separate document

12 ARE THE PARTICIPANTS COMPETENT TO GIVE INFORMED CONSENT ON THEIR OWN BEHALF? **Yes**
NOTE: Children and young adults under the age of 16 years (or 18 years if still at school) require parental/caregiver consent as do adults with disabilities that limit comprehension and consent. Such participants should be provided with a suitable information sheet and an assent form where practicable.

If No, please explain:

(a) Why they are not competent to give informed consent on their own behalf.

(b) How consent will be obtained.

D RISK, DECEPTION, PRIVACY

13. WHERE WILL THE PROJECT BE CONDUCTED?
NOTE: It is recommended that interviews be conducted in public spaces and where possible, not in private homes. In the case of research involving children, young adults and participants with disabilities, an adult other than the researcher is required to be present.

The experiments will be conducted where the participants are located. This will require the equipment to be brought to the organizations contacted for recruitment and potentially the homes of the participants.

14. FORESEEABLE RISKS TO THE PARTICIPANTS
If the answer to any of these questions is "Yes", please indicate briefly the nature of the risk and what actions you could take, or support mechanisms you could rely on, if a participant should become injured, distressed or offended while taking part in this project.
Support should not be undertaken by researcher. At the very least a list of support services should be included in the information sheet and also participants made

aware of the possibility in the information sheet.

(a) Is there any risk to physical well-being? **No**

If yes describe processes in place:

It is noted that there will not be an actual wheel chair involved, only the controller. Therefore there will not be a risk of driving the chair.

The controller will adhere to the safety standards followed by the University of Canterbury and will be checked by a technician.

(b) Could participation involve mental stress or emotional distress? **Yes**

If yes describe processes in place:

The ICF questions may be uncomfortable to answer for some participants, however they will not be asked any questions beyond what describing their daily life is like on a 0-4 scale. The ICF questionnaires are commonly used.

(c) Is there a possibility of giving moral or cultural offence? **No**

If Yes, describe processes in place and consultation/awareness undertaken:

15. IS DECEPTION INVOLVED AT ANY STAGE OF THE PROJECT? **No**

NOTE: The use in the information sheet or consent form or questionnaire of a title which differs from the project title given in this application form, in order not to reveal the real aim of the project, is considered to be a form of deception however mild.

If Yes, please

(a) Explain how and why it is to be used and how the participants will be 'debriefed' following their participation in the project.

(b) Attach a copy of the debriefing sheet prepared for use by the researcher or for distribution to the participants after their participation in the project or after the completion of the project.

16. WILL INFORMATION ABOUT THE SUBJECTS BE OBTAINED FROM THIRD PARTIES? **Yes**

This includes 'snowball' recruitment and also the accessing of potential participants via a third party.

In general third party contact information should not be given directly to the researcher – participants should contact the researcher and/or agree to be contacted.

If Yes, please state:

(a) The identity of the third party or parties.

CCS, Burwood Hospital and St. John of God

(b) Why such information is needed.

The third parties are required to obtain participants. They will be used to pass on invitations to potential participants. The participants will then be able to contact the researcher on their own accord.

Effectively the third parties will not be providing contact information to the researcher directly but will be a 'middle man' passing on the contact information of the researcher to the participants rather than passing the contact information of the participants to the researcher.

(c) Whether appropriate consents for access to such information have been or will be obtained.

It is at the discretion of the third parties to allow contact information to be passed onto the researcher or to pass on the request of the researcher to potential participants.

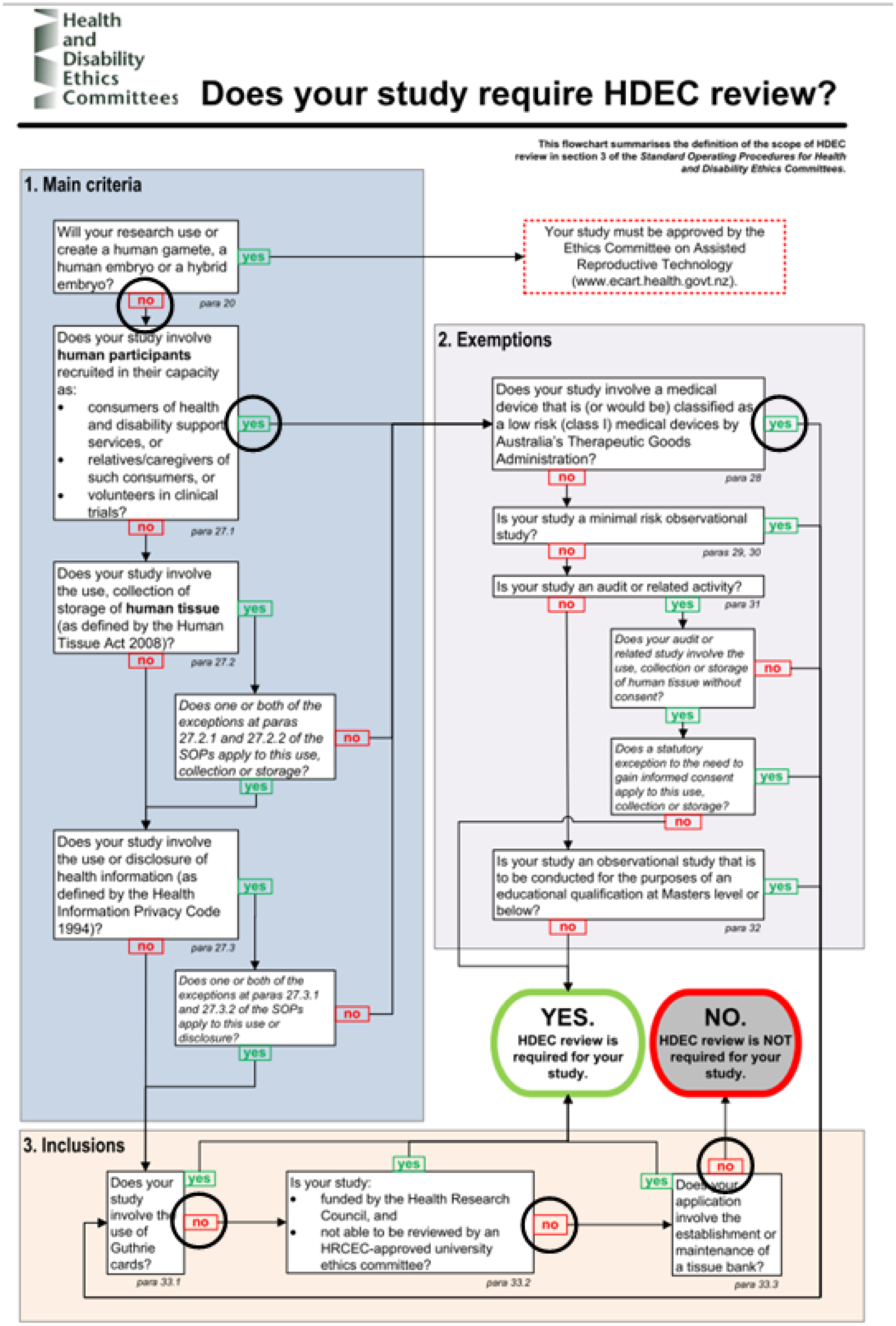
(d) Whether the use of such data in your research project needs the consent of the participants.

Only potential participants contact information will be sought i.e. name, email, phone number.

NOTE: It may happen that by virtue of your job, you have right of access to information concerning the participants. Such information may have been given by the participants for a particular purpose or collated by yourself or colleagues in the normal course of your job. The use of such information for a quite different purpose (i.e., a research project culminating in some form of report) may well require that potential participants at least be informed that their agreement to participate may involve such use. The Information Privacy Principles should be consulted for guidance in this area.

data will be included in the thesis and remaining raw data including participant contact details and video recordings will be deleted no more than 5 years after the project's completion.

A 5 Health and Disability Ethics Requirement Flow Chart



A 6 Participant Consent From and Information Sheet

See following pages

UC college of Engineering
T: +64 3 364 2608
E: collegeofengineering@canterbury.ac.nz

College of Engineering
University of Canterbury
Te Whare Wanaga o Waitaha
Private Bag 4800
Christchurch 8140
New Zealand

CONSENT FORM

Error Tracking on Wheelchair Interfaces

I have read and understood the description of the above-named project. On this basis I agree to participate as a participant in the project, and I consent to publication of the results of the project with the understanding that confidentiality will be preserved.

I understand that results may be published in a journal or other publication, as well as part of the Master's in Engineering thesis of Rory Horne and that this would be available publicly via the UC library database.

I understand that that my name or any personally identifying information will not be used in any part of the results, data, and final report.

I understand also that I may, at any time, withdraw from the project, including withdrawal of any information I have provided.

I understand that my participation is not part of credit or assessment for any course.

I note that this research has been reviewed and approved by the Department of Psychology, University of Canterbury.

I understand that I can contact Rory Horne or Dirk Pons regarding any concerns I may have regarding my participation in this study (contact details provided on the information sheet or through the University of Canterbury).

NAME (please print): …………………………………………………………….

Signature:

Date:

UC college of Engineering
T: +64 3 364 2608
E: collegeofengineering@canterbury.ac.nz

College of Engineering
University of Canterbury
Te Whare Wanaga o Waitaha
Private Bag 4800
Christchurch 8140
New Zealand
www.engf.canterbury.ac.nz

Researcher contact information
Contact email:
rory.horne@pg.canterbury.ac.nz
Contact phone: +64 03 364 2987 ext 8390

Human Ethics Committee,
University of Canterbury,
Private Bag 4800

Error Tracking on Wheelchair Interfaces

Information Sheet for Participant

My name is Rory Horne I am a postgraduate master's student in mechanical engineering. My thesis project is based on identifying errors made by users operating an electronic wheel chair. The objective of this experiment is to create a model describing human error. From this experiment I will identify what errors are made, the cause of the errors and ultimately use the information to confirm the model created. The hope is to improve controller design for electronic wheel chairs.

Your involvement in this project will require you to activate and use several functions on the electronic wheel chair controller. The entire process should take from 30 to 60 minutes. These tasks will be standard functions of the chair. How to perform these will be explained to you before you control the chair yourself. You will have several tempts to complete each task. A portion of the tasks will be done while you asked to reply to single words.

You will be given questionnaires to complete several times during the experiment between tasks.

The controller is attached to wheelchair that may be positioned next to you if it is inconvenient for you to sit in the test chair.

A camera attached to the chair will record your hand operating the controls and the controller inputs will be recorded electronically. The camera will only show the controls and your hand operating the control.

There will be no follow up to this investigation but if you have any further questions please feel free to contact the researcher (information available above).

You may receive a copy of the project results by contacting the researcher at the conclusion of the project; the project is scheduled to end at in the closing months of this year.

Participation is voluntary and you have the right to withdraw at any stage without penalty. If you withdraw, any information relating to you will be removed provided this request is submitted within one week of your participation (researcher contact information above). The recording of your participation will have a number associated with it that you are also given, if possible this number should be provided with your withdrawal request.

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public and if for any reason it is it will not be done so without your prior consent. To ensure anonymity and confidentiality the only recording of your person will be of your hand operating the controls (there will be no sound recordings kept). Video recording along with any other personal information gained (such as age, gender etc) will be destroyed by the end of 2018 and prior to this be kept in a secure computer file at the University of Canterbury.

Results of this investigation will be used in a thesis which is a public document and will be available through the UC Library. A copy of the thesis will be given directly to Dynamic Controls, the company which is involved in this project.

The project is being carried out as a part of masters in engineering thesis by Rory Horne (rory.horne@pg.canterbury.ac.nz) under the supervision of Dirk Pons (dirk.pons@canterbury.ac.nz) and Deak Helton (deak.helton@canterbury.ac.nz), who can be contacted at email address given. They will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

Complaints may be addressed to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch, Email: human-ethics@canterbury.ac.nz

If you agree to participate in the study, you are asked to complete the consent form and return Rory Horne before the testing begins.

Regards

Rory Horne

A 7 Organisations Contacted for Participant Recruitment

A 7.1 University of Canterbury Disability Services

General Enquiries to Room 214
Level 2, Puaka-James Hight building
University of Canterbury
Phone: +64 3 364 2350 or ext 6350
Email: disabilities@canterbury.ac.nz

A 7.2 The Laura Fergusson Trust Canterbury Inc.

279 Ilam Road
Ilam
Christchurch
8053

Phone: (03) 351 6047

A 7.3 Ministry of Health Disability Support Services

Phone: 0800 373 664

Email: disability@moh.govt.nz

A 7.4 St John of God Halswell

26 Nash Road
Halswell
Christchurch 8025

Tel: 03 338 2009
Email: enquiries.halswell@sjog.org.nz

A 8 Participant Questions before Experiment Start

A 8.1 Demographic Questions

Age:___

Gender:

- ☐ M
- ☐ F

What is your experience using an electronic wheelchair? (Please tick which is most appropriate)

- ☐ No experience
- ☐ Have briefly operated (10min or less) an electronic wheel chair 1 to 5 times
- ☐ Have briefly operated (10min or less) an electronic wheel chair more than 5 times
- ☐ Have used an electronic wheel chair for more than 1 day continually
- ☐ Have used an electronic wheelchair for more than 1 day
- ☐ Have used an electronic wheel for a significant portion of my life

When was did the experience indicated above occur? (Please tick which is most appropriate)

- ☐ More than 10 years ago
- ☐ More than 5 years ago
- ☐ More than 1 years ago
- ☐ In the last year
- ☐ In the last 3 months
- ☐ In the last week
- ☐ I have used an electronic wheel often for several years

How confident out of are you at operating new technology? Please circle indicate on the bar below

Not Confident at All

Very Confident

A 8.2 ICF

On the following page are several measures of body functions listed from the World Health Organisation's International Classification of Functioning.

Please give each of these measures a 0 to 4 rating for yourself corresponding to the ratings described below

0	No impairment, you have no difficulty
1	Mild impairment, a difficulty is present less than 25% of the time, with an intensity you can tolerate and which has happened rarely over the last 30 days.
2	Moderate impairment, a difficulty that is present less than 50% of the time, with an intensity, which is interfering in your day to day life and which has happened occasionally over the last 30 days.
3	Severe impairment, a difficulty is present more than 50% of the time, with intensity, which is partially disrupting to your day to day life and which has happened frequently over the last 30 days.
4	Complete impairment, a difficulty that is present more than 95% of the time, with an intensity, which is totally disrupting to your day to day life and which happens every day over the last 30 day

The Following Pages Contain the Questions…

Quality of psychomotor functions

At what level are your motor functions in performing sequenced movements such as typing? Please circle one

Text from ICF (B1471): Quality of psychomotor functions - Mental functions that produce nonverbal behaviour in the proper sequence and character of its subcomponents, such as hand and eye coordination, or gait.

0	1	2	3	4

Control of simple voluntary movements

At what level are your functions associated with control over and coordination of simple or isolated voluntary movement such as reaching to grab a glass? Please circle one

Text from ICF (B7600): Control of simple voluntary movements - Functions associated with control over and coordination of simple or isolated voluntary movements

0	1	2	3	4

Auditory perception

At what level are your abilities at discriminating sounds, tones, pitches and other noises? Please circle one

Text from ICF (B1560): Auditory perception - Mental functions involved in discriminating sounds, tones, pitches and other acoustic stimuli.

0	1	2	3	4

Visual perception

At what level are your abilities at discriminating shape, size and colour of objects and images? Please circle one

Text from ICF (B1561): Visual perception - Mental functions involved in discriminating shape, size, colour and other ocular stimuli.

0	1	2	3	4

Tactile perception

At what level are your abilities at feeling differences in texture by touch, such as rough or smooth? Please circle one

Text from ICF (B1564): Tactile perception - Mental functions involved in distinguishing differences in texture, such as rough or smooth stimuli, detected by touch.

0	1	2	3	4

Acquiring skills

At what level are your abilities to learn and develop basic sets of actions to perform a new skill or task, such as manipulating tools or playing games like chess? Please circle one

Text from ICF (D155): Acquiring skills - Developing basic and complex competencies in integrated sets of actions or tasks so as to initiate and follow through with the acquisition of a skill, such as manipulating tools or playing games like chess.

0	1	2	3	4

A 9 Participant Questions after each Task

A 9.1 Task Load Questionnaire

Mental Demand

How mentally demanding was the task?

Very Low Very High

Physical Demand

How physically demanding was the task?

Very Low Very High

Temporal Demand

How hurried or rushed did you feel performing the task?

Very Low Very High

Performance

How successful were you in accomplishing what you were asked to do?

Very Low Very High

Effort

How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed were you?

Very Low Very High

A 10 Participant Questions after Experiment

A 10.1 System Usability Scale

Please Circle a number below

I think that I would like to use this system frequently

Strongly Disagree	1	2	3	4	5	Strongly Agree

I found the system unnecessarily complex

Strongly Disagree	1	2	3	4	5	Strongly Agree

I thought the system was easy to use

Strongly Disagree	1	2	3	4	5	Strongly Agree

I think that I would need the support of a technical person to be able to use this system

Strongly Disagree	1	2	3	4	5	Strongly Agree

I found the various functions in this system were well integrated

Strongly Disagree	1	2	3	4	5	Strongly Agree

I thought there was too much inconsistency in this system

Strongly Disagree	1	2	3	4	5	Strongly Agree

I would imagine that most people would learn to use this system very quickly

Strongly Disagree	1	2	3	4	5	Strongly Agree

I found the system very cumbersome to use

Strongly Disagree	1	2	3	4	5	Strongly Agree

I felt very confident using the system

Strongly Disagree	1	2	3	4	5	Strongly Agree

I needed to learn a lot of things before I could get going with this system

Strongly Disagree	1	2	3	4	5	Strongly Agree

A 10.2 Usability Interview

General Comments

Do you have any general comments about how the system was to use?

Mistake

In the first attempt do you remember stopping and thinking about pressing a button that you thought was the correct one to press but turned out to be wrong? Can you remember when this was and what button?

What about in the other attempts?

Lapse

Did you feel like you were performing the task correctly during the any of the attempts but it turned out you pressed the wrong button? When did this happen?

Slip

During any of the tasks did you miss the button you intended to press?

A 11 Relevant sections of the World Health Organisation ICF

Websites used extensively

http://www.icfillustration.com/top_e.html

http://apps.who.int/classifications/icfbrowser/

A 11.1 Quality of psychomotor functions - B1471

Body Functions > Mental Functions > Specific Mental Functions > Psychomotor Functions > Quality of Psychomotor Functions

Mental functions that produce nonverbal behaviour in the proper sequence and character of its subcomponents, such as hand and eye coordination, or gait.

A 11.2 Control of simple voluntary movements - B7600

Body Functions > Neuromusculoskeletal and Movement-Related Functions > Movement Functions > Control of Voluntary Movement Functions > Control of simple voluntary movements

Functions associated with control over and coordination of simple or isolated voluntary movements

A 11.3 Auditory perception - B1560

Body Functions > Mental Functions > Specific Mental Functions > Perceptual Functions > Auditory Perception

Mental functions involved in discriminating sounds, tones, pitches and other acoustic stimuli.

A 11.4 Visual perception - B1561

Body Functions > Mental Functions > Specific Mental Functions > Perceptual Functions > Visual Perception

Mental functions involved in discriminating shape, size, colour and other ocular stimuli.

A 11.5 Tactile perception - B1564

Body Functions > Mental Functions > Specific Mental Functions > Perceptual Functions > Tactile Perception

Mental functions involved in distinguishing differences in texture, such as rough or smooth stimuli, detected by touch.

A 11.6 Acquiring skills - D155

Activities and Participation > Learning and Applying Knowledge > Basic Learning > Acquiring Skills

Developing basic and complex competencies in integrated sets of actions or tasks so as to initiate and follow through with the acquisition of a skill, such as manipulating tools or playing games like chess.

A 12 Experimental Procedure

1. Check equipment
    a. Test all functions
    b. Sync – can controller be reset i.e. turned on/off without needing to reset?
    c. Reset to start position
2. Greet participant (ask them to turn off/put on silent their phone)
3. Gain Consent
    a. Give info sheet and consent form, allow participant to read forms
    b. Ask participant if they have read and understand information sheet and consent form
    c. Have participant to sign consent form
4. Give first assessments
    a. Demographic assessments
    b. ICF assessments
5. Explain basic operations of controller and experiment to participant
    a. General process and objectives of experiment and tasks
    b. Which buttons they need to use
    c. What these buttons control and that the joystick activates the centre icon
6. Task 1 – turn on/off lights
    a. Explain objective of tasks and how they will know they have completed task
    b. Reset controller
    c. Ask participant to perform task
    d. Give TLX assessment
    e. Repeat from step b. until all four attempts are completed, only giving TLX assessment after the fourth attempt
7. Task 2 – raise/lower chair
    a. Explain objective of tasks and how they will know they have completed task
    b. Reset controller
    c. Ask participant to perform task
    d. Give TLX assessment
    e. Repeat from step b. until all four attempts are completed, only giving TLX assessment after the fourth attempt
8. Task 4 – enable/disable clock
    a. Explain objective of tasks and how they will know they have completed task
    b. Reset controller

c. Ask participant to perform task
d. Give TLX assessment
e. Repeat from step b. until all four attempts are completed, only giving TLX assessment after the fourth attempt

9. Cognitive loading
    a. Explain Cog. Loading word association task
    b. Explain task, that all previous tasks will need to be performed
    c. Reset controller
    d. Ask participants to perform all tasks in order
    e. Give TLX assessment
    f. Repeat from step b. until all four attempts are completed, only giving TLX assessment after the fourth attempt
10. Interview and Thank
    a. Conduct close out interview
    b. Remind them they can contact me for results
    c. Ask for any final comments on system
    d. Give reward
    e. Thank and lead out participant

A 13 Raw Data

This is the data collected directly from the participants with minimal processing.

See following pages for data

Note that participant 1 is not included as their results were discarded due to poor experimental procedure.

			Demo Q's						ICF Q's					
Participant	**Assistance**	**Poor Video**	**Age**	**Gender**		**Exp with chair**	**Time with chair (note 0 means NA)**	**Confi. (0-10 scale)**	**Psyc Motor**	**Movement**	**Audio**	**Visual**	**Tactile**	**Skills**
2			21	F	1	2	3	5	0	0	0	1	0	0
3			21	M	0	2	6	7	0	0	0	0	0	0
4			21	M	0	2	6	9	0	0	0	0	0	0
5			20	F	1	2	6	8	0	0	0	0	0	0
6			21	M	0	1	0	5	0	0	0	0	0	0
7			21	M	0	2	5	7	0	0	0	0	0	0
8			20	F	1	1	0	9	0	0	0	0	0	0
9			22	M	0	3	5	8	0	0	0	0	0	1
10	1		53	F	1	6	0	4	1	2	1	1	1	3
11			65	M	0	2	3	6	1	0	0	0	0	2
12	1		69	M	0	5	4	6	1	0	0	0	0	1
13			43	F	1	6	7	7	1	1	1	0	0	1
14	1		31	M	0	1	0	7	0	0	0	0	0	0
15	1		24	M	0	1	0	10	0	0	0	0	0	0
16	1		27	F	1	6	7	10	3	4	1	0	0	4
17	1		17	M	0	6	7	3	0	2	0	1	0	3
18			61	M	0	6	7	9	2	1	1	1	1	1
19	1		57	F	1	6	7	6	3	2	0	0	0	3
20	1		56	F	1	5	3	6	3	1	0	0	1	2
21	1		69	F	1	6	7	4	4	3	0	0	1	2
22	1		64	M	0	1	0	6	2	1	0	0	0	3
23	1	1	69	F	1	1	0	4	2	2	1	2	1	1
24	1	1	65	M	0	6	7	7	1	1	0	0	3	1
25	1	1	52	M	0	5	3	8	1	0	0	0	0	3
26		1	34	F	1	3	4	7	0	0	0	0	0	0
27	1	1	71	M	0	6	7	1	3	2	1	1	1	3
28	1	1	49	F	1	1	0	2	3	2	0	0	0	1
29		Y	20	M	0	1	0	7	1	0	0	0	0	0
30		1	20	F	1	1	0	7	0	0	0	0	0	0
31		1	19	F	1	1	0	7	0	0	0	0	0	0
32		1	21	F	1	1	0	8	0	0	0	0	0	0
33	1	1	59	M	0	1	0	6	0	0	0	0	0	0
34	1	1	57	F	1	1	0	7	0	0	0	0	0	0
35		1	53	F	1	1	0	6	0	0	0	0	0	0
36		1	54	M	0	1	0	7	0	0	0	0	0	0
37	1	1	51	F	1	1	0	4	0	0	0	0	0	0
38	1	1	65	M	0	2	2	3	0	0	1	1	0	0
39	1	1	61	F	1	2	2	5	0	0	1	1	0	1
40	1	1	56	M	0	1	0	4	1	0	0	1	0	1
41	1	1	52	M	0	1	0	6	0	0	0	0	0	0

Chair Task Attempt Times - Seconds

Participant	Task Start	Task End	1 [sec]	1[m:s]		2		2[m:s]		3		3[m:s]		4		4[m:s]
2	40	113	73	0	13	24	11	0	13	19	6	0	13	18	5	0
3	12	40	28	0	10	18	8	0	9	13	4	0	9	14	5	0
4	12	53	41	0	9	19	10	0	10	17	7	0	12	18	6	0
5	11	53	42	0	9	15	6	0	9	14	5	0	7	11	4	0
6	28	107	79	0	10	17	7	0	8	15	7	0	8	15	7	0
7	48	78	30	0	9	15	6	0	8	14	6	0	9	15	6	0
8	9	89	80	0	8	15	7	0	9	14	5	0	8	12	4	0
9	10	43	33	0	10	16	6	0	9	13	4	0	10	14	4	0
10	11	382	371	0	14	153	139	0	8	180	172	0	10	126	116	0
11	13	249	236	0	9	33	24	0	9	108	99	0	9	43	34	0
12	9	155	146	0	10	68	58	0	11	63	52	0	11	34	23	0
13	20	224	204	0	9	29	20	0	9	46	37	0	10	36	26	0
14	12	40	28	0	11	24	13	0	10	24	14	0	11	19	8	0
15	18	124	106	0	9	21	12	0	12	16	4	0	11	17	6	0
16	Participant awas unable to complete (attempt recorded)															
17	Participant awas unable to complete (attempt recorded)															
18	24	79	55	0	10	220	210	0	10	60	50	0	10	32	22	0
19	Participant was unable to to confidently attempt															
20	Participant was unable to to confidently attempt															
21	Participant was unable to to confidently attempt															
22	Participant was unable to to confidently attempt															
23	15	84	69	0	7	538	531	0	11	164	153	0	10	187	177	0
24	15	298	283	0	11	148	137	0	10	141	131	0	Participant chose to end due to frustration			
25	17	434	417	0	11	78	67	0	9	76	67	0	12	139	127	0
26	9	28	19	0	8	12	4	0	7	11	4	0	8	11	3	0
27	Participant was unable to to confidently attempt															
28	26	85	59	0	8	14	6	0	8	13	5	0	8	21	13	0
29	10	102	92	0	8	15	7	0	8	14	6	0	7	10	3	0
30	15	43	28	0	7	14	7	0	8	13	5	0	7	13	6	0
31	16	74	58	0	3	8	5	0	6	16	10	0	6	14	8	0
32	10	55	45	0	16	23	7	0	8	13	5	0	8	13	5	0
33	28	600	572	0	9	127	118	0	9	37	28	0	10	33	23	0
34	13	479	466	0	12	50	38	0	10	30	20	0	13	33	20	0
35	18	314	296	0	11	59	48	0	7	19	12	0	8	16	8	0
36	11	42	31	0	10	17	7	0	8	14	6	0	8	12	4	0
37	8	55	47	0	10	118	108	0	22	34	12	0	6	16	10	0
38	16	240	224	0	12	304	292	0	12	45	33	0	7	60	53	0
39	48	282	234	0	27	109	82	0	10	32	22	0	10	34	24	0
40	14	203	189	0	9	31	22	0	8	20	12	0	9	20	11	0
41	17	101	84	0	10	20	10	0	8	16	8	0	8	27	19	0

	TLX Chair 1 (TLX 0-10 scale)							TLX Chair 2						
Participant	Mental	Physical	Temporal	Performance	Effort	Frustration	Average	Mental	Physical	Temporal	Performance	Effort	Frustration	Average
2	9	1	9	7	11	9	8	3	1	2	19	2	2	5
3	10	2	8	13	14	6	9	2	2	2	18	3	2	5
4	7	1	7	17	5	12	8	3	0	5	19	4	2	6
5	2	1	1	19	2	1	4	0	0	1	19	1	1	4
6	4	0	0	18	5	0	5	0	0	0	19	0	1	3
7	2	0	1	19	0	0	4	0	0	0	19	1	0	3
8	11	0	10	0	4	16	7	0	0	0	19	0	0	3
9	8	2	13	19	10	8	10	1	0	1	19	2	1	4
10	5	1	2	13	7	6	6	15	5	6	7	13	10	9
11	17	3	7	15	18	7	11	6	1	1	16	4	4	5
12	3	0	0	14	2	0	3	2	0	0	15	0	1	3
13	15	1	2	8	12	6	7	2	1	1	19	2	1	4
14	6	0	0	20	2	1	5	2	0	0	20	1	0	4
15	14	0	14	10	14	20	12	0	0	0	0	0	0	0
16	Participant Was not able to complete tasks. Attempt at Ch1 recorded													
17	Participant was not able to complete tasks. Attempt at Ch1 recorded													
18	2	0	1	14	3	2	4	2	1	1	5	1	3	2
19	Participant was not able to complete tasks. Attempt at Ch2 recorded													
20	Participant was not able to complete tasks. Attempt at Ch1 recorded													
21	Participant was not able to complete tasks. Attempt at Ch1 recorded													
22	Participant chose to end experiment													
23	7	2	1	8	2	2	4	6	4	5	14	13	15	10
24	16	1	1	10	0	5	6	Participant chose to end task						
25	1	2	1	10	13	1	5	10	0	1	18	3	0	5
26	6	0	4	16	3	7	6	0	0	0	20	1	0	4
27	Participant was unable to complete tasks													
28	2	2	1	18	1	2	4	0	0	0	20	1	1	4
29	5	0	1	16	2	0	4	1	0	0	18	0	0	3
30	4	1	3	15	3	3	5	0	0	0	20	1	0	4
31	6	1	0	3	10	5	4	2	1	0	20	1	1	4
32	2	1	1	17	2	2	4	1	1	1	19	1	1	4
33	20	1	19	1	20	20	14	12	1	14	10	13	9	10
34	8	1	12	5	5	18	8	4	0	6	17	3	4	6
35	15	1	5	2	15	18	9	3	0	3	18	5	3	5
36	10	2	4	12	10	3	7	3	3	5	17	3	2	6
37	11	1	10	4	10	12	8	6	1	9	15	5	4	7
38	2	0	10	0	0	0	2	7	0	15	12	2	4	7
39	18	2	15	19	19	15	15	10	2	8	16	10	9	9
40	13	1	3	11	1	4	6	3	0	0	18	0	1	4
41	11	6	10	13	13	6	10	1	1	1	20	1	1	4

	Light Task Attempt Times - Seconds															
Participant	Task Start	Task End	1	1[m:s]		2		2[m:s]		3		3[m:s]		4		4[m:s]
2	18	74	56	0	13	18	5	0	9	14	5	0	9	14	5	0
3	9	15	6	0	9	12	3	0	9	11	2	0	9	11	2	0
4	14	23	9	0	10	16	6	0	9	13	4	0	9	13	4	0
5	8	20	12	0	8	15	7	0	8	12	4	0	12	16	4	0
6	9	20	11	0	9	15	6	0	9	12	3	0	12	15	3	0
7	10	16	6	0	10	14	4	0	12	15	3	0	9	12	3	0
8	8	14	6	0	8	12	4	0	8	11	3	0	8	11	3	0
9	10	17	7	0	10	15	5	0	8	11	3	0	10	13	3	0
10	11	56	45	0	11	209	198	0	17	187	170	0	12	161	149	0
11	9	141	132	0	9	21	12	0	9	15	6	0	9	12	3	0
12	9	65	56	0	11	28	17	0	10	19	9	0	10	17	7	0
13	10	85	75	0	10	24	14	0	9	20	11	0	9	19	10	0
14	9	20	11	0	11	15	4	0	10	15	5	0	9	14	5	0
15	13	34	21	0	10	15	5	0	13	20	7	0	10	14	4	0
16	Participant was unable to to confidently attempt															
17	Participant was unable to to confidently attempt															
18	25	141	116	0	9	45	36	0	28	73	45	0	11	26	15	0
19	Participant was unable to to confidently attempt															
20	Participant was unable to to confidently attempt															
21	Participant was unable to to confidently attempt															
22	Participant was unable to to confidently attempt															
23	28	175	147	0	9	73	64	0	9	29	20	0	10	17	7	0
24	Participant chose to end															
25	22	331	309	0	10	262	252	0	14	286	272	0	10	230	220	0
26	14	18	4	0	7	11	4	0	10	13	3	0	9	11	2	0
27	Participant was unable to to confidently attempt															
28	15	106	91	0	10	16	6	0	11	17	6	0	9	12	3	0
29	6	47	41	0	7	11	4	0	6	9	3	0	8	11	3	0
30	7	19	12	0	9	14	5	0	8	11	3	0	7	10	3	0
31	7	32	25	0	6	12	6	0	6	9	3	0	6	9	3	0
32	7	21	14	0	10	14	4	0	8	11	3	0	6	9	3	0
33	12	38	26	0	11	19	8	0	7	12	5	0	8	11	3	0
34	17	50	33	0	13	56	43	0	10	15	5	0	8	11	3	0
35	18	28	10	0	11	15	4	0	9	13	4	0	9	12	3	0
36	15	39	24	0	12	22	10	0	7	24	17	0	8	17	9	0
37	11	21	10	0	7	13	6	0	8	13	5	0	9	12	3	0
38	16	54	38	0	15	55	40	0	10	23	13	0	3	8	5	0
39	9	34	25	0	9	43	34	0	10	37	27	0	9	23	14	0
40	14	80	66	0	10	17	7	0	9	12	3	0	10	14	4	0
41	11	31	20	0	6	15	9	0	8	13	5	0	5	7	2	0

	TLX Lights 2							TLX Lights 2						
Participant	Mental	Physical	Temporal	Performance	Effort	Frustration	Average	Mental	Physical	Temporal	Performance	Effort	Frustration	Average
2	8	0	9	17	2	2	6	2	1	2	18	2	2	5
3	9	2	4	17	5	2	7	1	1	1	18	1	1	4
4	5	0	1	20	2	2	5	0	0	0	20	0	0	3
5	0	1	1	20	1	1	4	0	0	0	20	0	0	3
6	2	0	0	19	0	0	4	0	0	0	19	1	0	4
7	0	0	0	19	0	0	3	0	0	0	19	1	0	3
8	10	0	0	16	0	2	5	0	0	0	20	0	0	3
9	7	1	1	19	1	1	5	0	0	0	19	0	0	3
10	3	2	2	17	3	2	5	5	2	4	10	8	13	7
11	15	0	1	8	11	3	6	3	0	1	15	1	0	3
12	1	1	2	17	3	2	4	1	0	0	16	0	1	3
13	14	1	1	11	12	13	9	1	1	1	18	2	3	4
14	4	0	1	20	2	0	5	1	0	0	20	1	1	4
15	13	0	4	10	5	5	6	0	0	0	0	0	0	0
16														
17														
18	0	1	2	12	2	2	3	8	1	2	10	10	10	7
19														
20														
21														
22														
23	5	8	4	12	4	5	6	8	5	5	14	3	4	7
24														
25	0	1	2	19	1	1	4	1	2	2	19	1	1	4
26	5	0	0	19	0	0	4	0	0	0	20	0	0	3
27														
28	1	1	1	13	1	1	3	1	0	0	17	0	0	3
29	4	1	1	15	5	0	4	1	0	0	17	0	0	3
30	3	1	2	16	2	2	4	0	1	1	20	1	0	4
31	4	1	0	17	1	2	4	1	1	1	20	1	1	4
32	1	2	1	19	2	2	5	1	1	1	20	1	1	4
33	18	1	6	17	11	4	10	4	0	3	18	6	2	6
34	6	2	4	12	4	5	6	2	1	2	19	0	1	4
35	13	0	1	20	2	1	6	1	0	1	20	1	1	4
36	8	2	5	12	10	5	7	8	2	7	12	12	8	8
37	10	1	4	13	5	3	6	3	1	3	14	4	2	4
38	0	0	3	3	0	0	1	0	0	0	0	0	0	0
39	16	0	8	3	5	5	6	4	0	3	0	3	2	2
40	11	1	4	8	1	4	5	1	0	1	18	2	1	4
41	9	4	3	17	4	3	7	0	0	0	20	0	0	3

Participant	Clock Task Attempt Times - Seconds															
	Task Start	Task End	1	1[m:s]		2		2[m:s]		3		3[m:s]		4		4[m:s]
2	21	208	187	0	11	30	19	0	6	24	18	0	6	17	11	0
3	9	120	111	0	9	30	21	0	10	22	12	0	9	19	10	0
4	9	128	119	0	9	26	17	0	9	23	14	0	9	22	13	0
5	10	42	32	0	10	26	16	0	9	19	10	0	9	19	10	0
6	9	83	74	0	8	25	17	0	9	19	10	0	9	18	9	0
7	9	307	298	0	9	26	17	0	9	30	21	0	8	27	19	0
8	12	264	252	0	9	70	61	0	8	35	27	0	9	28	19	0
9	9	42	33	0	8	65	57	0	8	24	16	0	10	21	11	0
10	15	642	627	0	14	548	534	0	14	329	315	0	12	215	203	0
11	15	626	611	0	7	386	379	0	10	118	108	0	8	151	143	0
12	11	474	463	0	12	194	182	0	15	80	65	0	10	70	60	0
13	8	326	318	0	11	137	126	0	18	304	286	0	9	107	98	0
14	16	320	304	0	11	27	16	0	11	26	15	0	10	20	10	0
15	10	400	390	0	10	113	103	0	4	29	25	0	1	22	21	0
16	Participant was unable to to confidently attempt															
17	Participant was unable to to confidently attempt															
18	10	700	690	0	10	311	301	0	10	102	92	0	9	45	36	0
19	Participant was unable to to confidently attempt															
20	Participant was unable to to confidently attempt															
21	Participant was unable to to confidently attempt															
22	Participant was unable to to confidently attempt															
23	Participant was unable to complete task															
24	Participant chose to end															
25	Participant chose to end															
26	10	112	102	0	9	24	15	0	10	26	16	0	9	28	19	0
27	Participant was unable to complete task															
28	9	328	319	0	9	172	163	0	Participant chose to end							
29	9	44	35	0	8	26	18	0	9	17	8	0	9	17	8	0
30	10	189	179	0	8	20	12	0	8	16	8	0	7	14	7	0
31	12	224	212	0	12	68	56	0	8	20	12	0	6	16	10	0
32	7	138	131	0	6	89	83	0	7	36	29	0	7	22	15	0
33	10	422	412	0	15	211	196	0	11	105	94	0	11	36	25	0
34	10	393	383	0	9	61	52	0	10	46	36	0	8	38	30	0
35	20	88	68	0	11	54	43	0	11	66	55	0	10	41	31	0
36	11	339	328	0	9	78	69	0	9	59	50	0	9	35	26	0
37	7	307	300	0	8	300	292	0	9	121	112	0	9	45	36	0
38	11	822	811	0	11	175	164	0	12	108	96	0	11	238	227	0
39	14	289	275	0	11	147	136	0	10	84	74	0	10	72	63	0
40	15	568	553	0	8	76	68	0	7	55	48	0	8	65	57	0
41	10	401	391	0	10	176	166	0	7	48	41	0	11	30	19	0

	TLX Clock 1							TLX Clock 2						
Participant	Mental	Physical	Temporal	Performance	Effort	Frustration	Average	Mental	Physical	Temporal	Performance	Effort	Frustration	Average
2	15	1	14	5	15	15	11	4	1	3	18	4	3	6
3	12	3	4	11	12	2	7	9	4	3	16	6	2	7
4	12	0	6	10	10	12	8	4	0	2	20	5	7	6
5	3	2	0	18	3	1	5	1	0	0	20	2	0	4
6	1	0	0	18	2	0	4	0	0	0	19	0	0	3
7	14	0	4	14	14	10	9	2	0	1	17	1	1	4
8	18	0	17	0	18	19	12	13	1	4	9	13	17	10
9	12	3	13	17	13	10	11	13	3	2	18	10	11	10
10	14	3	14	10	3	14	10	13	3	10	10	4	10	8
11	18	7	0	8	6	5	7	11	0	0	10	14	1	6
12	18	3	4	10	4	4	7	3	1	5	0	7	5	4
13	15	1	11	5	15	16	11	17	2	1	16	17	17	12
14	16	0	2	8	10	5	7	3	0	0	20	1	0	4
15	20	0	0	0	20	20	10	2	0	0	20	20	0	7
16														
17														
18	1	1	1	8	3	1	3	1	1	1	11	2	1	3
19														
20														
21														
22														
23	participant attempted but not completed													
24														
25	Participant was not able to continue tasks													
26	10	0	3	16	5	10	7	5	1	2	18	2	4	5
27														
28	15	0	0	3	3	6	5	Participant chose not to continue. Becam frustrated. Clock attempt 2 is recorded						
29	7	1	0	17	6	0	5	4	3	1	15	4	4	5
30	14	1	3	9	9	8	7	3	1	1	20	3	0	5
31	10	2	5	7	3	16	7	1	1	1	19	1	1	4
32	8	3	2	10	8	12	7	1	1	1	20	1	1	4
33	20	4	19	1	20	20	14	19	2	15	3	14	16	12
34	7	3	15	6	10	18	10	7	1	4	16	6	7	7
35	10	0	5	10	10	3	6	10	0	3	17	10	2	7
36	18	2	18	3	18	18	13	14	2	10	8	14	14	10
37	14	2	15	4	15	16	11	12	1	9	10	11	10	9
38	15	0	18	0	18	19	12	10	2	12	6	10	10	8
39	20	0	14	20	20	10	14	13	0	15	16	13	16	12
40	19	0	14	1	7	19	10	14	0	1	13	4	7	7
41	19	17	17	7	18	16	16	6	7	5	15	10	6	8

	Cog Task Attempt Times - Seconds															
Participant	ideo Start [se	ideo End [se	1 [sec]	1 [m:s]		2		2[m:s]		3		3 [m:s]		4		4 [m:s]
2	17	83	66	0	15	57	42	0	14	58	44	0	14	55	41	0
3	14	79	65	0	13	64	51	0	13	48	35	0	15	56	41	0
4	15	73	58	0	18	57	39	0	14	47	33	0	17	43	26	0
5	10	305	295	0	10	77	67	0	9	67	58	0	15	64	49	0
6	14	67	53	0	12	64	52	0	15	58	43	0	12	42	30	0
7	12	64	52	0	14	79	65	0	15	63	48	0	12	67	55	0
8	16	241	225	0	19	96	77	0	22	85	63	0	16	75	59	0
9	12	63	51	0	9	48	39	0	10	39	29	0	11	39	28	0
10	Participant Chose to Stop															
11	9	482	473	0	11	359	348	0	Participant Chose to Stop							
12	11	164	153	0	10	180	170	0	11	98	87	0	10	82	72	0
13	29	628	599	0	Participant Chose to Stop											
14	10	73	63	0	10	64	54	0	17	58	41	0	11	40	29	0
15	10	85	75	0	12	62	50	0	9	54	45	0	11	49	38	0
16	Participant was unable to to confidently attempt															
17	Participant was unable to to confidently attempt															
18	14	436	422	0	22	450	428	0	24	219	195	0	25	349	324	0
19	Participant was unable to to confidently attempt															
20	Participant was unable to to confidently attempt															
21	Participant was unable to to confidently attempt															
22	Participant was unable to to confidently attempt															
23	Participant was unable to to confidently attempt															
24																
25																
26	9	72	63	0	9	55	46	0	9	61	52	0	9	46	37	0
27			0	0			0	0			0	0			0	0
28																
29	8	35	27	0	8	29	21	0	9	30	21	0	8	33	25	0
30	9	58	49	0	8	52	44	0	8	39	31	0	8	39	31	0
31	7	83	76	0	7	64	57	0	6	40	34	0	8	41	33	0
32	10	66	56	0	16	64	48	0	10	40	30	0	9	40	31	0
33	8	134	126	0	21	118	97	0	9	180	171	0	9	145	136	0
34	8	129	121	0	14	104	90	0	18	92	74	0	11	58	47	0
35	15	135	120	0	15	105	90	0	9	82	73	0	8	69	61	0
36	10	118	108	0	9	91	82	0	14	93	79	0	9	81	72	0
37	9	169	160	0	11	86	75	0	9	89	80	0	12	71	59	0
38	15	308	293	0	11	259	248	0	17	163	146	0	14	121	107	0
39	30	555	525	0	14	200	186	0	13	117	104	0	13	82	69	0
40	7	461	454	0	15	164	149	0	10	234	224	0	16	113	97	0
41	13	117	104	0	10	60	50	0	9	69	60	0	9	49	40	0

							Cog. TLX 2						
Participant	Physical	Temporal	Performance	Effort	Frustration	Average	Mental	Physical	Temporal	'erformanc	Effort	Frustration	Average
2	1	11	14	9	9	8	7	1	7	17	8	6	8
3	4	8	11	13	4	7	11	3	6	13	13	4	8
4							6	0	10	14	13	8	9
5	3	6	16	9	5	7	3	1	3	19	2	1	5
6	0	9	17	8	0	7	4	0	4	18	3	0	5
7	0	4	4	5	1	4	5	0	6	6	11	2	5
8	0	19	0	19	19	11	19	0	19	13	19	13	14
9	3	13	17	14	10	11	13	3	12	14	14	6	10
10													
11	0	1	9	18	1	7	15	5	0	5	16	0	7
12	4	5	11	17	5	9	18	4	7	12	17	5	11
13	1	10	1	19	19	11							
14	0	5	20	5	1	8	4	0	2	20	2	1	5
15	0	0	15	20	0	8	5	0	0	20	0	0	4
16													
17													
18	1	10	10	10	5	9	12	2	3	10	10	7	7
19													
20													
21													
22													
23													
24													
25													
26	1	7	12	6	7	10	9	2	8	12	6	8	8
27													
28													
29	1	4	14	10	1	10	6	0	4	14	6	2	5
30	1	4	16	11	7	12	10	1	5	19	9	4	8
31	3	7	10	5	9	11	6	1	4	12	2	4	5
32	2	1	10	8	10	11	6	1	1	10	2	4	4
33	3	19	4	18	19	16	19	2	18	4	17	18	13
34	3	17	10	16	19	17	11	3	6	11	6	9	8
35	0	15	7	19	10	14	12	0	7	12	10	5	8
36	2	14	10	14	16	15	12	2	12	12	14	14	11
37	2	16	7	16	16	16	12	2	10	10	10	11	9
38	3	15	3	13	14	14	4	2	8	9	5	10	6
39	0	14	10	17	14	16	10	0	10	10	10	7	8
40	1	18	2	17	17	16	17	2	13	11	13	13	12
41	17	19	12	19	19	21	11	15	9	16	13	11	13

	SUS								
Participant	Freq. +	Complex -		Easy +	Support -		Integration +	Inconsistency -	
2	4	2	4	2	2	4	2	4	2
3	4	2	4	3	1	5	4	3	3
4	2	5	1	2	2	4	3	2	4
5	4	2	4	3	1	5	4	3	3
6	3	1	5	4	1	5	2	1	5
7	2	3	3	2	1	5	3	5	1
8	1	5	1	1	5	1	1	5	1
9	1	4	2	2	1	5	2	4	2
10	1	4	2	4	3	3	3	3	3
11	2	5	1	3	2	4	5	2	4
12	4	4	2	4	1	5	4	4	2
13	1	4	2	2	5	1	1	4	2
14	1	4	2	1	3	3	3	3	3
15	1	3	3	2	1	5	4	5	1
16									
17									
18	1	3	3	1	3	3	3	4	2
19									
20									
21									
22									
23	4	2	4	2	2	4	3	2	4
24	1	5	1	1	5	1	3	5	1
25	4	3	3	2	5	1	5	5	1
26	4	4	2	2	2	4	3	3	3
27									
28	2	4	2	1	4	2	2	4	2
29	3	1	5	4	1	5	3	1	5
30	3	2	4	4	1	5	3	1	5
31	1	2	4	3	1	5	5	1	5
32	2	4	2	2	2	4	1	4	2
33	1	5	1	1	5	1	1	5	1
34	2	5	1	3	5	1	1	4	2
35	3	3	3	3	2	4	3	2	4
36	3	4	2	3	4	2	2	4	2
37	2	4	2	3	4	2	2	4	2
38	1	5	1	4	4	2	4	4	2
39	1	5	1	1	5	1	1	5	1
40	2	4	2	2	3	3	2	3	3
41	1	5	1	1	4	2	2	4	2

A 14 Condensed Raw Data

This is the raw data seen in Appendix 12 processed for use in the statistical analysis

See following pages for data

Note that participant 1 is not included as their results were discarded due to poor experimental procedure.

Participant	Group		Assistance	Demo Q's Age	Gender	Exp	Time (note 0 means NA)	Confidence	Total ICF	Average ICF	Any ICF > 3	Stopped after 3 Tasks	Not Completed
2	1	<35		21	1	2	3	5	0	0.00	0	0	0
3	1	<35		21	0	2	6	7	0	0.00	0	0	0
4	1	<35		21	0	2	6	9	0	0.00	0	0	0
5	1	<35		20	1	2	6	8	0	0.00	0	0	0
6	1	<35		21	0	1	0	5	0	0.00	0	0	0
7	1	<35		21	0	2	5	7	0	0.00	0	0	0
8	1	<35		20	1	1	0	9	0	0.00	0	0	0
9	1	<35		22	0	3	5	8	1	0.33	0	0	0
10	3	DNC	1	53	1	6	0	4	6	2.00	1	0	1
11	2	50+		65	0	2	3	6	3	1.00	0	0	1
12	2	50+	1	69	0	5	4	6	2	0.67	0	0	0
13	3	DNC		43	1	6	7	7	3	1.00	0	0	1
14	1	<35	1	31	0	1	0	7	0	0.00	0	0	0
15	1	<35	1	24	0	1	0	10	0	0.00	0	0	0
16	3	DNC	1	27	1	6	7	10	11	3.67	1	1	1
17	3	DNC	1	17	0	6	7	3	5	1.67	1	1	1
18	2	50+		61	0	6	7	9	4	1.33	0	0	0
19	3	DNC	1	57	1	6	7	6	8	2.67	1	1	1
20	3	DNC	1	56	1	5	3	6	6	2.00	1	1	1
21	3	DNC	1	69	1	6	7	4	9	3.00	1	1	1
22	3	DNC	1	64	0	1	0	6	6	2.00	1	1	1
23	3	DNC	1	69	1	1	0	4	5	1.67	0	1	1
24	3	DNC	1	65	0	6	7	7	3	1.00	0	1	1
25	3	DNC	1	52	0	5	3	8	4	1.33	1	1	1
26	1	<35		34	1	3	4	7	0	0.00	0	0	0
27	3	DNC	1	71	0	6	7	1	8	2.67	1	1	1
28	3	DNC	1	49	1	1	0	2	6	2.00	1	1	1
29	1	<35		20	0	1	0	7	1	0.33	0	0	0
30	1	<35		20	1	1	0	7	0	0.00	0	0	0
31	1	<35		19	1	1	0	7	0	0.00	0	0	0
32	1	<35		21	1	1	0	8	0	0.00	0	0	0
33	2	50+	1	59	0	1	0	6	0	0.00	0	0	0
34	2	50+	1	57	1	1	0	7	0	0.00	0	0	0
35	2	50+		53	1	1	0	6	0	0.00	0	0	0
36	2	50+		54	0	1	0	7	0	0.00	0	0	0
37	2	50+	1	51	1	1	0	4	0	0.00	0	0	0
38	2	50+	1	65	0	2	2	3	0	0.00	0	0	0
39	2	50+	1	61	1	2	2	5	1	0.33	0	0	0
40	2	50+	1	56	0	1	0	4	2	0.67	0	0	0
41	2	50+	1	52	0	1	0	6	0	0.00	0	0	0

Participant	Group	Chair Attempt 1	Chair Attempt 2	Chair Attempt 3	Chair Attempt 4	Chair TLX1 Ave Demand	Chair TLX1 Performance	Chair TLX4 Ave Demand	Chair TLX4 Performance
2	1	73	11	6	5	9.5	7	2.25	19
3	1	28	8	4	5	9.5	13	2.25	18
4	1	41	10	7	6	7.75	17	3.5	19
5	1	42	6	5	4	1.5	19	0.75	19
6	1	79	7	7	7	2.25	18	0.25	19
7	1	30	6	6	6	0.75	19	0.25	19
8	1	80	7	5	4	10.25	0	0	19
9	1	33	6	4	4	9.75	19	1.75	18
10	3	371	139	172	116	5	13	11	7
11	2	236	24	99	34	12.25	15	3.75	16
12	2	146	58	52	23	1.25	14	0.75	15
13	3	204	20	37	26	8.75	8	1.5	19
14	1	28	13	14	8	2.25	20	0.75	20
15	1	106	12	4	6	15.5	10	0	0
16	3								
17	3								
18	2	55	210	50	22	2	14	1.75	5
19	3								
20	3								
21	3								
22	3								
23	3	69	531	153	177	3	8	9.75	14
24	3	283	137	131					
25	3	417	67	67	127	4	10	3.5	18
26	1	19	4	4	3	5	16	0.25	20
27	3								
28	3	59	6	5	13	1.5	18	0.5	20
29	1	92	7	6	3	2	16	0.25	18
30	1	28	7	5	6	3.25	15	0.25	20
31	1	58	5	10	8	5.25	3	1	20
32	1	45	7	5	5	1.75	17	1	19
33	2	572	118	28	23	19.75	1	12	10
34	2	466	38	20	20	10.75	5	4.25	17
35	2	296	48	12	8	13.25	2	3.5	18
36	2	31	7	6	4	6.75	12	3.25	17
37	2	47	108	12	10	10.7	4	6	15
38	2	224	292	33	53	3	0	7	12
39	2	234	82	22	24	16.75	19	9.25	16
40	2	189	22	12	11	5.25	11	1	18
41	2	84	10	8	19	10	13	1	20

Participant	Group	Light Attempt 1	Light Attempt 2	Light Attempt 3	Light Attempt 4	Lights TLX1 Ave Demand	Lights TLX1 Performance	Lights TLX4 Ave Demand	Lights TLX4 Performance
2	1	56	5	5	5	5.25	17	2.00	18
3	1	6	3	2	2	5.00	17	1.00	18
4	1	9	6	4	4	2.50	20	0.00	20
5	1	12	7	4	4	0.75	20	0.00	20
6	1	11	6	3	3	0.50	19	0.00	20
7	1	6	4	3	3	0.00	19	0.25	19
8	1	6	4	3	3	3.00	16	0.00	20
9	1	7	5	3	3	2.50	19	0.00	19
10	3	45	198	170	149	2.50	17	7.50	10
11	2	132	12	6	3	7.50	8	1.25	15
12	2	56	17	9	7	2.25	11	0.50	16
13	3	75	14	11	10	10.00	11	1.75	18
14	1	11	4	5	5	1.75	20	0.75	20
15	1	21	5	7	4	6.75	10	0.00	0
16	3								
17	3								
18	2	116	36	45	15	1.50	12	7.50	10
19	3								
20	3								
21	3								
22	3								
23	3	147	64	20	7	4.50	12	5.00	14
24	3								
25	3	309	252	272	220	1.00	19	1.25	19
26	1	4	4	3	2	1.25	19	0.00	20
27	3								
28	3	91	6	6	3	1.00	13	0.25	17
29	1	41	4	3	3	2.50	15	0.25	17
30	1	12	5	3	3	2.25	16	0.50	20
31	1	25	6	3	3	1.75	17	1.00	20
32	1	14	4	3	3	1.50	19	1.00	20
33	2	26	8	5	3	9.75	17	3.75	18
34	2	33	43	5	3	4.75	12	1.25	19
35	2	10	4	4	3	4.25	20	1.00	20
36	2	24	10	17	9	7.00	12	8.75	12
37	2	10	6	5	3	5.55	13	2.95	14
38	2	38	40	13	5	0.75	3	0.00	0
39	2	25	34	27	14	8.50	3	3.00	0
40	2	66	7	3	4	5.00	8	1.25	18
41	2	20	9	5	2	4.75	17	0.00	20

Participant	Group	Clock Attempt 1	Clock Attempt 2	Clock Attempt 3	Clock Attempt 4	Clcok TLX1 Ave Demand	Clock TLX1 Performance	Clcok TLX4 Ave Demand	Clock TLX4 Performance
2	1	187	19	18	11	14.75	5	3.5	18
3	1	111	21	12	10	7.5	11	5	16
4	1	119	17	14	13	10	10	4.5	20
5	1	32	16	10	10	1.75	18	0.75	20
6	1	74	17	10	9	0.75	18	0	19
7	1	298	17	21	19	10.5	14	1.25	17
8	1	252	61	27	19	18	0	11.75	9
9	1	33	57	16	11	12	17	9	18
10	3	627	534	315	203	11.25	10	9.25	10
11	2	611	379	108	143	7.25	8	6.5	10
12	2	463	182	65	60	7.5	10	5	0
13	3	318	126	286	98	14.25	5	13	16
14	1	304	16	15	10	8.25	8	1	20
15	1	390	103	25	21	15	0	5.5	20
16	3								
17	3								
18	2	690	301	92	36	1.5	8	1.25	11
19	3								
20	3								
21	3								
22	3								
23	3								
24	3								
25	3								
26	1	102	15	16	19	7	16	3.25	18
27	3								
28	3	319	163						
29	1	35	18	8	8	3.25	17	3.25	15
30	1	179	12	8	7	8.5	9	1.75	20
31	1	212	56	12	10	8.5	7	1	19
32	1	131	83	29	15	7.5	10	1	20
33	2	412	196	94	25	19.75	1	16	3
34	2	383	52	36	30	12.5	6	6	16
35	2	68	43	55	31	7	10	6.25	17
36	2	328	69	50	26	18	3	13	8
37	2	300	292	112	36	14.95	4	10.35	10
38	2	811	164	96	227	17.5	0	10.5	6
39	2	275	136	74	63	16	20	14.25	16
40	2	553	68	48	57	14.75	1	6.5	13
41	2	391	166	41	19	17.5	7	6.75	15

Participant	Group	Cog Attempt 1	Cog Attempt 2	Cog Attempt 3	Cog Attempt 4	Cog TLX1 Ave Demand	Cog TLX1 Performance	Cog TLX4 Ave Demand	Cog TLX4 Performance
2	1	66	42	44	41	10.250	14	7	17
3	1	65	51	35	41	9.500	11	8.5	13
4	1	58	39	33	26	10.036	11	9.25	14
5	1	295	67	58	49	7.500	16	2.25	19
6	1	53	52	43	30	6.250	17	2.75	18
7	1	52	65	48	55	4.750	4	6	6
8	1	225	77	63	59	19.000	0	17.5	13
9	1	51	39	29	28	13.000	17	11.25	14
10	3								
11	2	473	348			9.250	9		
12	2	153	170	87	72	11.250	11	11.75	12
13	3	599				16.750	1		
14	1	63	54	41	29	5.250	20	2.25	20
15	1	75	50	45	38	9.000	15	1.25	20
16	3								
17	3								
18	2	422	428	195	324	8.750	10	8	10
19	3								
20	3								
21	3								
22	3								
23	3								
24	3								
25	3								
26	1	63	46	52	37	7.500	12	7.75	12
27	3								
28	3								
29	1	27	21	21	25	6.750	14	4.5	14
30	1	49	44	31	31	8.750	16	7	19
31	1	76	57	34	33	7.500	10	4	12
32	1	56	48	30	31	5.750	10	3.25	10
33	2	126	97	171	136	18.750	4	18	4
34	2	121	90	74	47	17.000	10	8	11
35	2	120	90	73	61	15.250	7	8.5	12
36	2	108	82	79	72	14.500	10	13	12
37	2	160	75	80	59	16.050	7	10.85	10
38	2	293	248	146	107	14.750	3	6.75	9
39	2	525	186	104	69	13.750	10	9.25	10
40	2	454	149	224	97	17.250	2	14	11
41	2	104	50	60	40	18.750	12	11	16

Participant	Group	SUS Average Rating
2	1	2.7
3	1	3.9
4	1	2.5
5	1	3.7
6	1	4
7	1	2.6
8	1	1
9	1	2.3
10	3	2.6
11	2	3.1
12	2	3.6
13	3	1.4
14	1	2
15	1	2.6
16	3	
17	3	
18	2	2.5
19	3	
20	3	
21	3	
22	3	
23	3	3.3
24	3	1.8
25	3	3.3
26	1	2.7
27	3	
28	3	2.2
29	1	4.3
30	1	4.3
31	1	4.1
32	1	2.5
33	2	1.4
34	2	1.8
35	2	3.3
36	2	2.7
37	2	2.24
38	2	2
39	2	1
40	2	2.4
41	2	1.4

A 15 Raw Interview Data

This is the raw data from the recorded interview notes categorised into themes presented in the report.

See following pages for data.

Note that participant 1 is not included as their results were discarded due to poor experimental procedure.

Participants	General Comments	Physical Layout	Icon Clarity	M	L	S
2	to many lists on side and bottom	Long distance from stick to buttons	hard to distinguish at first but know when learnt	numerous on 1st attempt	No	No
3	clock should Hve simple toggle some options/features were hidden first time hard eady after that No menu 'themes'	link between button and screen not clear for I&II	some confusing icons but mostly ok	numerous on 1st attempt	No	No
4	didn't like controller controller menu was dissililar to other device menus Menu layers didn't link	relation between buttons and screen not clear	Icons generally clear but light icon not obvious	numerous on 1st attempt	No	No
5	hard to start, easy when learnt	fine	two clock were confusing	guessing on Ch and L thining on C	few time during cog load	No
6	hard to start, easy to once learnt	fine	some icons not initially intuitive	some on first attempts	No	No
7	no 'back' button function of button/joystick not always consistant/clear joystick in particular	Joystick uncomfortble distance from buttons	brightness icon (sunshine) often confused with lights many icons not obvious	actions generally intuitive (not conciuous thought) apart from clock	missed some intermediate steps	No
8	Not logical Could use text not icons No flow in tasks Couldn't get 'out' of tasks, no back button Very poor experience overall Didn't know if had completed cog task (many following did this too)	Not great	light icons were confusing to many 'clocks'	Yes, esp clocks	Yes	No
9	Didn't intuitivly know where to go in menus, got lost	seemed to have to many buttons	mostly good	when setting clock moslty went by gut feel	maybe on chair	No
10	was uncomfortable using it as not a computer person unexpected results from buttons became a little frustated hard to get out of wrong menus, hard to learn paths	physically it was usable Sometimes buttons were not pressed correctly	some icons were confused, such as brightness and light			
11	liked the looped menus	fine	fine	Yes, esp clocks	yes	maybe
12	lots of featuers	fine	fine for good vision screen had some refection	some, Particulry cog	some, particulary cog	maybe
13	Didn't realise bottom menu had several options for some featuers e.g. "not sure how I got X on the clock" MY OBSERVATION: many don't get the correct mental model of the vertical and horizontal menu interaction The clock on/off icon is confusing as it doesn't say if the clock is on or off Navigation is confusing Touch screen would be good Big change in complexity between tasks Ilogical flow between tasks Some icons hard to see in corners	Good Screen a bit small	Results of actions not always clear			
14	Not intuitive No 'cursor'. Don't know whats selected link between physical button and action on screen not obvious	fine seems to have a lot of keys	clocks confusing MY OBERVATION: there are a lot of icons that could represent lights e.g. screen brightness, dark/light screen change			
15	confusion with link between vertical and horizontal menues a lot of 'scrolling' required					
16	NA					
17	NA					

18	Confusion with link between vertical and horizontal menues Felt it was more of a memory challenge to complete the tasks than it was understanding the layout and icons Trouble understanding navigagtion, link between screen and buttons Paticipant completed Cog2 at 6:34 but didn realise the clock was turned on, thought task was done at 7:30	trouble operating buttons, tended to hold them down				
23	observation: tried to use as a touch screen multiple times observation: had to remind to use red/blue buttons multiple times observation: participant used joystick for movement a great deal over buttons Assistance given					
24	Note: participant had a prothetic but was still able to operate the controller Participant chose to end the experiment out of frustration Quote by participant (P24 Ch3 2:20-2:26) "I can'r be bothered with it, honestly, I'm starting to get confused."					
25	Tried to use as touch screen. Could identify sought icon but had trouble 'getting to it using the buttons.					
28	had a lot of trouble making the conecting of scrolling to find the tick and cross in the clock task					
29	used joystick left/right	fine	ok	maybe once on task1	maybe	No
31	did an action to turn off clock (pull joystick back) which was wong and became confused as whether clock was turned off or not					
32	didn't know how clock was turned off					
33	not user friendly menu navigation was difficult No prompts, no 'asking' want you want My Observations: icons on right hand menu don't link with icons that come up in centre when moved onto namely P icon People often get close to turning clock on/off i.e. Just need to push joystick forward then do the wrong thing. Common clock error is to pull joystick backwards.	fine	Not good. Seat colour for on and off hard to distingish i.e. Blue and geen hard to tell apart	Yes	Yes	No
34	menus didn't make sense, very inconsistant apart from clock once learnt was easier was scared to go beyond what was learnt	fine	fine	Yes	Yes	No
35	took a while to undersatnd meaning of blue/green on chiars	fine	fine	Yes	Yes	No
36	Clock was not intuitive	angle to operate button from joystick was poor. Bad ergonomics	fine	Yes	YEs	No
37	clock is to complex, to many steps	ok	ok	Yes	Yes	No
38	not user friendly would be worse if disabled	ok		yes	yes	no
40	difficult to know what was 'highlighted' regardign the seats (seems to be a common problem) Difficulty with the clock seems to break down the built mental model e.g. P40 could no longer find the orange clock after struggeling witht the clock task navigation is not that clear. Movement of icons on menus confusing	ok, joystick a little high Screen could be bigger	ok	Y	Yes	No
41	screen to small and hard to see	joystick stopped easy use of buttons	not great			

References

A brief history of ergonomics and user centered design. (2014). Retrieved August 24, 2014, from http://www.ergonomics.org.uk/iehf/a-brief-history/

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*(2), 219–245. doi:10.1037/0033-2909.121.2.219

Albert, W., & Dixon, E. (2003). Is this what you expected? The use of expectation measures in usability testing. Presented at the Usability Professionals Association, Scottsdale.

Allik, J., & Realo, A. (1997). Intelligence, academic abilities, and personality. *Personality and Individual Differences*, *23*(5), 809–814. doi:10.1016/S0191-8869(97)00103-7

Almanji, A., Claire Davies, T., & Susan Stott, N. (2014). Using cursor measures to investigate the effects of impairment severity on cursor control for youths with cerebral palsy. *International Journal of Human-Computer Studies*, *72*(3), 349–357. doi:10.1016/j.ijhcs.2013.12.003

Arnold, H. J., & Feldman, D. C. (1981). Social Desirability Response Bias in Self-Report Choice Situations. *Academy of Management Journal*, *24*(2), 377–385. doi:10.2307/255848

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. In Kenneth W. Spence and Janet Taylor Spence (Ed.), *Psychology of Learning and Motivation* (Vol. Volume 2, pp. 89–195). Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S0079742108604223

Austin, E. J., Deary, I. J., Whiteman, M. C., Fowkes, F. G. ., Pedersen, N. L., Rabbitt, P., … McInnes, L. (2002). Relationships between ability and personality:does intelligence contribute positively to personal and social adjustment? *Personality and Individual Differences*, *32*(8), 1391–1411. doi:10.1016/S0191-8869(01)00129-5

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 556–559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews. Neuroscience*, *4*(10), 829–39. doi:http://dx.doi.org.ezproxy.canterbury.ac.nz/10.1038/nrn1201

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. doi:10.1016/S0022-5371(75)80045-4

Balsamo, M., Lauriola, M., & Saggino, A. (2012). Personality and college major choice: which come first? *Psychology (Irvine)*, *3*(5), 399+.

Bandura, A. (1977). *Social learning theory* (Vol. viii). Oxford, England: Prentice-Hall.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology*, *63*(3), 575–582. doi:10.1037/h0045925

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. doi:10.1080/10447310802205776

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280–289. doi:10.1016/j.tics.2007.05.005

Barrick, M. R., & Mount, M. K. (1991). THE BIG FIVE PERSONALITY DIMENSIONS AND JOB PERFORMANCE: A META-ANALYSIS. *Personnel Psychology*, *44*(1), 1–26. doi:10.1111/j.1744-6570.1991.tb00688.x

Bastian, A. J. (2006). Learning to predict the future: the cerebellum adapts feedforward movement control. *Current Opinion in Neurobiology*, *16*(6), 645–649. doi:10.1016/j.conb.2006.08.016

Beecher, F. (2014). Ending the UX Designer Drought. Retrieved August 24, 2014, from http://boxesandarrows.com/ending-the-ux-designer-drought/

Beggs, W. D. A., Graham, J. C., Monk, T. H., Shaw, M. R. W., & Howarth, C. I. (1972). Can Hick's law and Fitts' law be combined? *Acta Psychologica*, *36*(5), 348–357. doi:10.1016/0001-6918(72)90031-5

Beh, H. C., Roberts, R. D., & Prichard-Levy, A. (1994). The relationship between intelligence and choice reaction time within the framework of an extended model of Hick's law: A preliminary report. *Personality and Individual Differences*, *16*(6), 891–897. doi:10.1016/0191-8869(94)90233-X

Bennett, S., Maton, K., & Kervin, L. (2008). The "digital natives" debate: A critical review of the evidence. *British Journal of Educational Technology*, *39*(5), 775–786. doi:10.1111/j.1467-8535.2007.00793.x

Bhushan, N., & Shadmehr, R. (1999). Computational nature of human adaptive control during learning of reaching movements in force fields. *Biological Cybernetics*, *81*(1), 39–60. doi:10.1007/s004220050543

Bickenbach, J. E., Chatterji, S., Badley, E. ., & Üstün, T. . (1999). Models of disablement, universalism and the international classification of impairments, disabilities and handicaps. *Social Science & Medicine*, *48*(9), 1173–1187. doi:10.1016/S0277-9536(98)00441-9

Blakemore, S., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself ? *NeuroReport*, *11*(11), pp.R11–R16.

Bødker, S. (2006). When Second Wave HCI Meets Third Wave Challenges. In *Proceedings of the 4th Nordic Conference on Human-computer Interaction: Changing Roles* (pp. 1–8). New York, NY, USA: ACM. doi:10.1145/1182475.1182476

Boff, K. R. (2006). Revolutions and shifting paradigms in human factors & ergonomics. *Applied Ergonomics*, *37*(4), 391–399. doi:10.1016/j.apergo.2006.04.003

Bonsiepe, G. (1968). A method of quantifying order in typographic design. *Journal of Typographic Research*, *2*, 203–220.

Boorse, C. (1975). On the Distinction Between Disease and Illness. *Philosophy and Public Affairs*, *5*(1), 49–68.

Boorse, C. (1977). Health as a Theoretical Concept. *Philosophy of Science*, *44*(4), 542–573.

Boren, T., & Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *IEEE Transactions on Professional Communication*, *43*(3), 261–278. doi:10.1109/47.867942

Borsci, S., Federici, S., & Lauriola, M. (2009). On the dimensionality of the System Usability Scale: a test of alternative measurement models. *Cognitive Processing*, *10*(3), 193–197. doi:10.1007/s10339-009-0268-9

Brand, C. R. (1994). Open to experience—closed to intelligence: Why the “Big Five” are really the “Comprehensive Six.” *European Journal of Personality*, *8*(4), 299–310. doi:10.1002/per.2410080407

Brisenden, S. (1986). Independent Living and the Medical Model of Disability. *Disability, Handicap & Society*, *1*(2), 173–178. doi:10.1080/02674648666780171

Broadbent, D. (1958). The selective nature of learning. In *Perception and communication* (pp. 244–267). Elmsford, NY, US: Pergamon Press.

Brooke, J. (1996). SUS-A quick and dirty usability scale. In *Usability Evaluation In Industry* (pp. 189–194). CRC Press.

Brooke, J. (2013). SUS: a retrospective. *Journal of Usability Studies 8(2)*, 29–40.

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct Measurement of Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 53–61. doi:10.1207/S15326985EP3801_7

Bruyère, S. M., Van Looy, S. A., & Peterson, D. B. (2005). The International Classification of Functioning, Disability and Health: Contemporary Literature Overview. *Rehabilitation Psychology*, *50*(2), 113–121. doi:10.1037/0090-5550.50.2.113

Bubic, A., Cramon, D. Y. von, & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4(25). doi:10.3389/fnhum.2010.00025

Burke, S. N. B., Carol A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*, *7*(1), 30–40. doi:10.1038/nrn1809

Butz, M. V., & Pezzulo, G. (2008). Benefits of Anticipations in Cognitive Agents. In G. Pezzulo, M. V. Butz, C. Castelfranchi, & R. Falcone (Eds.), *The Challenge of Anticipation* (pp. 45–62). Springer Berlin Heidelberg. Retrieved from http://link.springer.com.ezproxy.canterbury.ac.nz/chapter/10.1007/978-3-540-87702-8_3

Carlson, N. R., Heth, D., Miller, H., Donahoe, J., Martin, G. N., Carlson, N. R., … Martin, G. N. (2009). *Psychology: the science of behavior.* Harlow: Pearson. Retrieved from https://eprints.mdx.ac.uk/5654/

Carrier, L. M., Cheever, N. A., Rosen, L. D., Benitez, S., & Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of Americans. *Computers in Human Behavior*, *25*(2), 483–489. doi:10.1016/j.chb.2008.10.012

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3), 104–110. doi:10.1016/j.tics.2005.01.011

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, *44*(7), 1596–1603. doi:10.1016/j.paid.2008.01.003

Chamorro-Premuzic, T., Moutafi, J., & Furnham, A. (2005). The relationship between personality traits, subjectively-assessed and fluid intelligence. *Personality and Individual Differences*, *38*(7), 1517–1528. doi:10.1016/j.paid.2004.09.018

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332.

Chapuis, O., Blanch, R., & Beaudouin-Lafon, M. (2007). Fitts' Law in the Wild: A Field Study of Aimed Movements. Retrieved from http://hal.archives-ouvertes.fr/hal-00612026

Cheriton, D. R. (1976). Man-machine interface design for timesharing systems. In *Proceedings of the 1976 annual conference* (pp. 362–366). New York, NY, USA: ACM. doi:10.1145/800191.805617

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 213–218). New York, NY, USA: ACM. doi:10.1145/57167.57203

Chung, J., & Monroe, G. S. (2003). Exploring Social Desirability Bias. *Journal of Business Ethics*, *44*(4), 291–302.

Church, K. W., & Hanks, P. (1990). Word Association Norms, Mutual Information, and Lexicography. *Comput. Linguist.*, *16*(1), 22–29.

Comber, T. (2010). An analysis of entropy, redundancy and complexity in graphical user interfaces. *MBA Southern Cross University, Lismore, NSW*. Retrieved from http://epubs.scu.edu.au/theses/124

Comber, T., & Maltby, J. (1996). Investigating layout complexity. *Design, Specification, and Verification of Interactive Systems '96: Proceedings of the Eurographics Workshop, Springer, Wien; New York*. Retrieved from http://epubs.scu.edu.au/comm_pubs/3

Comber, T., & Maltby, J. (1997). Layout complexity: does it measure usability? *Human-Computer Interaction: Interact '97, International Conference on Human-Computer Interaction*. Retrieved from http://epubs.scu.edu.au/comm_pubs/2

Comber, T., & Maltby, J. R. (1995). Evaluating usability of screen designs with layout complexity. *School of Commerce and Management Papers*, 1.

Corbetta, M., & Shulman, G. L. (2002). Control of Goal-Directed and Stimulus-Driven Attention in the Brain. *Nature Reviews. Neuroscience*, *3*(3), 201–15. doi:http://dx.doi.org.ezproxy.canterbury.ac.nz/10.1038/nrn755

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*, 323–338. doi:10.1016/S0079-6123(07)00020-9

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. MIT Press.

Crews, D. E., & Zavotka, S. (2006). Aging, Disability, and Frailty: Implications for Universal Design. *Journal of PHYSIOLOGICAL ANTHROPOLOGY*, *25*(1), 113–118. doi:10.2114/jpa2.25.113

Davis, L. J. (2013). *The Disability Studies Reader*. Routledge.

De Feyter, T., Caers, R., Vigna, C., & Berings, D. (2012). Unraveling the impact of the Big Five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learning and Individual Differences*, *22*(4), 439–448. doi:10.1016/j.lindif.2012.03.013

Definition and Domains of ergonomics | IEA Website. (2014). Retrieved August 24, 2014, from http://www.iea.cc/whats/

Dekker, P. S., Cook, D. R. I., Johannesen, D. L., Sarter, D. N., & Woods, P. D. D. (2010). *Behind Human Error*. Ashgate Publishing, Ltd.

De Kleijn-De Vrankrijker, M. W. (2003). The long way from the International Classification of Impairments, Disabilities and Handicaps (ICIDH) to the International Classification of Functioning, Disability and Health (ICF). *Disability & Rehabilitation*, *25*(11-12), 561–564. doi:10.1080/09638280110110879

De Lange, M. A., & van Knippenberg, A. (2009). To err is human: How regulatory focus and action orientation predict performance following errors. *Journal of Experimental Social Psychology*, *45*(6), 1192–1199. doi:10.1016/j.jesp.2009.07.009

Demetriou, A., Kyriakides, L., & Avraamidou, C. (2003). The missing link in the relations between intelligence and personality. *Journal of Research in Personality*, *37*(6), 547–581. doi:10.1016/S0092-6566(03)00048-5

Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, *41*(1), 417–440. doi:10.1146/annurev.ps.41.020190.002221

Downton, A. (1991). *Engineering The Human-Computer Interface*. London.

Embrey, D. E. (1986). SHERPA: A systematic human error reduction and prediction approach. Retrieved from http://inis.iaea.org/Search/search.aspx?orig_q=RN:18074340

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716. doi:10.1038/35094565

Engel, G. L. (1977). The Need for a New Medical Model: A Challenge for Biomedicine. *Science*, *196*(4286), 129–136. doi:10.2307/1743658

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 97–120. doi:10.1016/j.trf.2005.04.012

Erk, S., Abler, B., & Walter, H. (2006). Cognitive modulation of emotion anticipation. *European Journal of Neuroscience*, *24*(4), 1227–1236. doi:10.1111/j.1460-9568.2006.04976.x

Escera, C., Alho, K., Schröger, E., & Winkler, I. (2000). Involuntary Attention and Distractibility as Evaluated with Event-Related Brain Potentials. *Audiology & Neurotology*, *5*(3-4), 151–66.

Federici, S., Micangeli, A., Ruspantini, I., Borgianni, S., Corradi, F., Pasqualotto, E., & Olivetti Belardinelli, M. (2005). Checking an integrated model of web accessibility and usability evaluation for disabled people. *Disability and Rehabilitation*, *27*(13), 781–790. doi:10.1080/09638280400014766

Ferner, R. E., & Aronson, J. K. (2006). Clarification of Terminology in Medication Errors: Definitions and Classification. *Drug Safety*, *29*(11), 1011.

Festinger, L., & Canon, L. K. (1965). Information about spatial location based on knowledge about efference. *Psychological Review*, *72*(5), 373–384. doi:10.1037/h0022205

Fisher, R. J., & Katz, J. E. (2000). Social-Desirability Bias and the Validity of Self-Reported Values. *Psychology & Marketing*, *17*(2), 105–120.

Fitts, P. M., & Peterson, J. R. (1964). Information Capacity of Discrete Motor Responses. *Journal of Experimental Psychology*, *67*, 103–112.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A Description of Think Aloud Method and Protocol Analysis. *Qualitative Health Research*, *3*(4), 430–441. doi:10.1177/104973239300300403

Freed, L., & Ishida, S. (1995). *History of Computers*. Hightstown, NJ, USA: Ziff-Davis Publishing Co.

Fridin, M., & Belokopytov, M. (2014). Embodied robot vs. virtual agent: involvement of pre-school children in motor task performance. *International Journal of Human-Computer Interaction*, doi:10.1080/10447318.2014.888500

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, *7*(3), 385–400. doi:10.1016/0191-8869(86)90014-0

Furnham, A., Forde, L., & Cotter, T. (1998). Personality and intelligence. *Personality and Individual Differences*, *24*(2), 187–192. doi:10.1016/S0191-8869(97)00169-4

Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and Intelligence: Gender, the Big Five, Self-Estimated and Psychometric Intelligence. *International Journal of Selection and Assessment*, *13*(1), 11–24. doi:10.1111/j.0965-075X.2005.00296.x

Furnham, A.-P. (2004). Estimating one's own personality and intelligence scores. *British Journal of Psychology*, *95*(2), 149–160.

Futrell, G. (1985). What Do You Look for in a New Engineer? *Journal of Management in Engineering*, *1*(1), 20–27. doi:10.1061/(ASCE)9742-597X(1985)1:1(20)

Gallistel, C. R. (1981). The organization of action: A new synthesis. *Behavioral and Brain Sciences*, *4*(04), 609–619. doi:10.1017/S0140525X00000480

Ghez, C., Gordon, J., & Ghilardi, M. F. (1995). Impairments of reaching movements in patients without proprioception. II. Effects of visual information on accuracy. *Journal of Neurophysiology*, *73*(1), 361–372.

Ghez, C., Gordon, J., Ghilardi, M. F., Christakos, C. N., & Cooper, S. E. (1990). Roles of Proprioceptive Input in the Programming of Arm Trajectories. *Cold Spring Harbor Symposia on Quantitative Biology*, *55*, 837–847. doi:10.1101/SQB.1990.055.01.079

Gibson, J. J. (1986). *The Ecological Approach To Visual Perception*. Psychology Press.

Gilbert, D. T., & Wilson, T. D. (2009). Why the brain talks to itself: sources of error in emotional prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1335–1341. doi:10.1098/rstb.2008.0305

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67. doi:10.1037/0033-295X.91.1.1

Glavin, R. J. (2010). Drug errors: consequences, mechanisms, and avoidance. *British Journal of Anaesthesia*, *105*(1), 76–82. doi:10.1093/bja/aeq131

Goodell, K. H., Cao, C. G. L., & Schwaitzberg, S. D. (2006). Effects of Cognitive Distraction on Performance of Laparoscopic Surgical Tasks. *Journal of Laparoendoscopic & Advanced Surgical Techniques*, *16*(2), 94–98. doi:10.1089/lap.2006.16.94

Goodley, D. (2001). "Learning Difficulties", the Social Model of Disability and Impairment: challenging epistemologies. *Disability & Society*, *16*(2), 207–231. doi:10.1080/09687590120035816

Gosling, S. D., Rentfrow, P. J., & Swann Jr., W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. doi:10.1016/S0092-6566(03)00046-1

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*(3), 377–96; discussion 396–442.

Harrison, R., Tomblen, D. T., & Jackson, T. A. (1955). Profile of the Mechanical Engineer III. Personality. *Personnel Psychology*, *8*(4), 469–490. doi:10.1111/j.1744-6570.1955.tb01224.x

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908. doi:10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Peter A. Hancock and Najmedin Meshkati (Ed.), *Advances in Psychology* (Vol. Volume 52, pp. 139–183). North-Holland. Retrieved from http://www.sciencedirect.com/science/article/pii/S0166411508623869

Hasan, Z. (1992). Role of proprioceptors in neural control. *Current Opinion in Neurobiology*, *2*(6), 824–829. doi:10.1016/0959-4388(92)90140-G

Hassenzahl, M. (2008). User Experience (UX): Towards an Experiential Perspective on Product Quality. In *Proceedings of the 20th International Conference of the Association Francophone D'Interaction Homme-Machine* (pp. 11–15). New York, NY, USA: ACM. doi:10.1145/1512714.1512717

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, *25*(2), 91–97. doi:10.1080/01449290500330331

Hazrati-Viari, A., Rad, A. T., & Torabi, S. S. (2012). The effect of personality traits on academic performance: The mediating role of academic motivation. *Procedia - Social and Behavioral Sciences*, *32*(0), 367–371. doi:10.1016/j.sbspro.2012.01.055

Helander, M. (2005). *A Guide to Human Factors and Ergonomics, Second Edition* (2nd ed.). CRC Press.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), 11–26. doi:10.1080/17470215208416600

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method (CREAM)*. Elsevier.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107–129. doi:10.1016/0001-6918(67)90011-X

Jaeger, P. T. (2009). LIFELONG INTERACTIONS: Persons with Disabilities and Intergenerational Universal Usability. *Interactions*, *16*(3), 66–67. doi:10.1145/1516016.1516031

James, W. (1890). *The Principles of Psychology* - (Vol. 2). LULU Press.

Jaspers, M. W. M., Steen, T., Bos, C. van den, & Geenen, M. (2004). The think aloud method: a guide to user interface design. *International Journal of Medical Informatics*, *73*(11–12), 781–795. doi:10.1016/j.ijmedinf.2004.08.003

Johansson, R. S., & Westling, G. (1984). Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research*, *56*(3), 550–564. doi:10.1007/BF00237997

Johnson, C. (1999). Why human error modeling has failed to help systems development. *Interacting with Computers*, *11*(5), 517–524. doi:10.1016/S0953-5438(98)00041-1

Johnson, G. I., Clegg, C. W., & Ravden, S. J. (1989). Towards a practical method of user interface evaluation. *Applied Ergonomics*, *20*(4), 255–260. doi:10.1016/0003-6870(89)90187-7

Jordan, M. I., & Rumelhart, D. E. (1992). Forward Models: Supervised Learning with a Distal Teacher. *Cognitive Science*, *16*(3), 307–354. doi:10.1207/s15516709cog1603_1

Judd, T., & Kennedy, G. (2011). Measurement and evidence of computer-based task switching and multitasking by “Net Generation” students. *Computers & Education*, *56*(3), 625–631. doi:10.1016/j.compedu.2010.10.004

Juul, J. (2012). *A Casual Revolution: Reinventing Video Games and Their Players*. MIT Press.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kappe, R., & van der Flier, H. (2010). Using multiple and specific criteria to assess the predictive validity of the Big Five personality factors on academic performance. *Journal of Research in Personality*, *44*(1), 142–145. doi:10.1016/j.jrp.2009.11.002

Karray, F., Alemzadeh, M., Saleh, J. A., & Arab, M. N. (2008). *Human-Computer Interaction: Overview on State of the Art*.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727. doi:10.1016/S0959-4388(99)00028-8

Kawato, M., Furukawa, K., & Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, *57*(3), 169–185. doi:10.1007/BF00364149

Kay, R. H. (1993a). An exploration of theoretical and practical foundations for assessing attitudes toward computers: The Computer Attitude Measure (CAM). *Computers in Human Behavior*, *9*(4), 371–386. doi:10.1016/0747-5632(93)90029-R

Kay, R. H. (1993b). A practical research tool for assessing ability to use computers: The Computer Ability Survey (CAS). *Journal of Research on Computing in Education*, *26*(1), 16.

Keates, S., Clarkson, P. J., Harrison, L.-A., & Robinson, P. (2000). Towards a Practical Inclusive Design Approach. In *Proceedings on the 2000 Conference on Universal Usability* (pp. 45–52). New York, NY, USA: ACM. doi:10.1145/355460.355471

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, *70*(6, Pt.1), 387–403. doi:10.1037/h0026739

Kichuk, S. L., & Wiesner, W. H. (1997). The big five personality factors and team performance: implications for selecting successful product design teams. *Journal of Engineering and Technology Management*, *14*(3–4), 195–221. doi:10.1016/S0923-4748(97)00010-6

Kirakowski, J., & Corbett, M. (1993). SUMI: the Software Usability Measurement Inventory. *British Journal of Educational Technology*, *24*(3), 210–212. doi:10.1111/j.1467-8535.1993.tb00076.x

Kirwan, B. (1992a). Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied Ergonomics*, *23*(5), 299–318. doi:10.1016/0003-6870(92)90292-4

Kirwan, B. (1992b). Human error identification in human reliability assessment. Part 2: Detailed comparison of techniques. *Applied Ergonomics*, *23*(6), 371–381. doi:10.1016/0003-6870(92)90368-6

Kirwan, B. (1994). *A Guide To Practical Human Reliability Assessment*. CRC Press.

Kirwan, B. (1998a). Human error identification techniques for risk assessment of high risk systems—Part 1: review and evaluation of techniques. *Applied Ergonomics*, *29*(3), 157–177. doi:10.1016/S0003-6870(98)00010-6

Kirwan, B. (1998b). Human error identification techniques for risk assessment of high risk systems—Part 2: towards a framework approach. *Applied Ergonomics*, *29*(5), 299–318. doi:10.1016/S0003-6870(98)00011-8

Kirwan, B., & Ainsworth, L. K. (1992). *A Guide To Task Analysis: The Task Analysis Working Group*. CRC Press.

Kline, P., & Lapham, S. L. (1992). Personality and faculty in British universities. *Personality and Individual Differences*, *13*(7), 855–857. doi:10.1016/0191-8869(92)90061-S

Komarraju, M., & Karau, S. J. (2005). The relationship between the big five personality traits and academic motivation. *Personality and Individual Differences*, *39*(3), 557–567. doi:10.1016/j.paid.2005.02.013

Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, *19*(1), 47–52. doi:10.1016/j.lindif.2008.07.001

Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(48), 13232–13240. doi:10.1523/JNEUROSCI.3481-07.2007

LaBerge, D. (1995). *Attentional Processing: The Brain's Art of Mindfulness*. Harvard University Press.

Lam, L. T. (2002). Distractions and the risk of car crash injury: The effect of drivers' age. *Journal of Safety Research*, *33*(3), 411–419. doi:10.1016/S0022-4375(02)00034-8

Lewis, J. R. (1991). Psychometric Evaluation of an After-scenario Questionnaire for Computer Usability Studies: The ASQ. *SIGCHI Bull.*, *23*(1), 78–81. doi:10.1145/122672.122692

Lewis, J. R. (1992). Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *36*(16), 1259–1260. doi:10.1177/154193129203601617

Lewis, J. R. (1994). Sample Sizes for Usability Studies: Additional Considerations. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *36*(2), 368–378. doi:10.1177/001872089403600215

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*(1), 57–78. doi:10.1080/10447319509526110

Lewis, J. R. (2002). Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction*, *14*(3-4), 463–488. doi:10.1080/10447318.2002.9669130

Lewis, J. R., & Sauro, J. (2009). The Factor Structure of the System Usability Scale. In M. Kurosu (Ed.), *Human Centered Design* (pp. 94–103). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-02806-9_12

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. *Accident Analysis & Prevention*, *42*(3), 881–890. doi:10.1016/j.aap.2009.05.001

Li, D. (2011, July). *Product instructions in the digital age* (phd). University of Leeds. Retrieved from http://etheses.whiterose.ac.uk/2110/

Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: a five-factor model perspective. *Medical Education*, *36*(11), 1050–1056. doi:10.1046/j.1365-2923.2002.01328.x

Lindley, R. H., Bathurst, K., Smith, W. R., & Wilson, S. M. (1993). Hick's law, IQ, and singularity or specificity of mind: A psychometric analysis. *Personality and Individual Differences*, *15*(2), 129–135. doi:10.1016/0191-8869(93)90019-Y

Llewellyn, A., & Hogan, K. (2000). The Use and Abuse of Models of Disability. *Disability & Society*, *15*(1), 157–165. doi:10.1080/09687590025829

Llinás, R. R. (2002). *I of the Vortex: From Neurons to Self*. MIT Press.

Lopez, R., Love, P., Edwards, D., & Davis, P. (2010). Design Error Classification, Causation, and Prevention in Construction Engineering. *Journal of Performance of Constructed Facilities*, *24*(4), 399–408. doi:10.1061/(ASCE)CF.1943-5509.0000116

Lucas, D. (2001, February). Human error prediction and controls: demonstrations made in COMAH safety cases. In *Proceedings of an IBC Conference on Human Error, London, 27th-28th February*.

Lutz, S., & Huitt, W. (2003). Information processing and memory: Theory and applications. *Educational Psychology Interactive*.

Maltby, J. (2007). Operational Complexity of Direct Manipulation Tasks in a Windows Environment. *Australasian Journal of Information Systems*, *2*(2). doi:10.3127/ajis.v2i2.402

Marseguerra, M., Zio, E., & Librizzi, M. (2006). Quantitative developments in the cognitive reliability and error analysis method (CREAM) for the assessment of human performance. *Annals of Nuclear Energy*, *33*(10), 894–910. doi:10.1016/j.anucene.2006.05.003

Maslow, A. H. (1970). *Motivation and personality* (3rd ed.). Harper & Row.

Mattox, E., Andersson. (2012). Strategies for Improving Patient Safety: Linking Task Type to Error Type. *Critical Care Nurse*, *32*(1), 52–78. doi:10.4037/ccn2012303

Mazzoni, P., & Krakauer, J. W. (2006). An Implicit Plan Overrides an Explicit Strategy during Visuomotor Adaptation. *The Journal of Neuroscience*, *26*(14), 3642–3645. doi:10.1523/JNEUROSCI.5317-05.2006

McGee, M. (2003). Usability Magnitude Estimation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(4), 691–695. doi:10.1177/154193120304700406

McGee, M. (2004). Master Usability Scaling: Magnitude Estimation and Master Scaling Applied to Usability Measurement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 335–342). New York, NY, USA: ACM. doi:10.1145/985692.985735

Mcmillan, S. J., & Morrison, M. (2006). Coming of age with the internet A qualitative exploration of how the internet has become an integral part of young people's lives. *New Media & Society*, *8*(1), 73–95. doi:10.1177/1461444806059871

McSweeney, R. (1992). *Sumi–a psychometric approach to software evaluation* (Unpublished MA (Qual) thesis in Applied Psychology). University College Cork, Cork, Ireland.

Meyer, D. E., & Kornblum, S. (1993). *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience*. MIT Press.

Miall, R. C., Haggard, P. N., & Cole, J. D. (1995). Evidence of a limited visuo-motor memory used in programming wrist movements. *Experimental Brain Research*, *107*(2), 267–280. doi:10.1007/BF00230047

Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Networks*, *9*(8), 1265–1279. doi:10.1016/S0893-6080(96)00035-4

Militello, L. G., & Hoffman, R. R. (2008). The Forgotten History of Cognitive Task Analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52*(4), 383–387. doi:10.1177/154193120805200439

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. doi:10.1037/h0043158

Mortenson, W. B., Miller, W. C., & Auger, C. (2008). Issues for the Selection of Wheelchair-Specific Activity and Participation Outcome Measures: A Review. *Archives of Physical Medicine and Rehabilitation*, *89*(6), 1177–1186. doi:10.1016/j.apmr.2008.01.010

Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: a study using the Neo Personality Inventory and the Myers–Briggs Type Indicator. *European Journal of Personality*, *17*(1), 79–94. doi:10.1002/per.471

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences*, *16*(1), 31–42. doi:10.1016/j.lindif.2005.06.003

Moutafi, J., Furnham, A., & Paltiel, L. (2004). Why is Conscientiousness negatively correlated with intelligence? *Personality and Individual Differences*, *37*(5), 1013–1022. doi:10.1016/j.paid.2003.11.010

Moutafi, J., Furnham, A., & Paltiel, L. (2005). Can personality factors predict intelligence? *Personality and Individual Differences*, *38*(5), 1021–1033. doi:10.1016/j.paid.2004.06.023

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience Letters*, *98*(2), 217–221. doi:10.1016/0304-3940(89)90513-2

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280. doi:10.1002/ejsp.2420150303

Neisser, U. (1967). *Cognitive psychology*. East Norwalk, CT, US: Appleton-Century-Crofts.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology*. W. H. Freeman.

Newell, A. F., & Gregor, P. (2002). Design for older and disabled people – where do we go from here? *Universal Access in the Information Society*, *2*(1), 3–7. doi:10.1007/s10209-002-0031-9

Ngo, D. C. L., Teo, L. S., & Byrne, J. G. (2000). Formalising guidelines for the design of screen layouts. *Displays*, *21*(1), 3–15. doi:10.1016/S0141-9382(00)00026-3

Nickerson, R. (1969). Man-Computer Interaction: A Challenge for Human Factors Research. *IEEE Transactions on Man Machine Systems*, *10*(4), 164–180. doi:10.1109/TMMS.1969.299924

Nielsen, J. (1993). Iterative user-interface design. *Computer*, *26*(11), 32–41. doi:10.1109/2.241424

Nielsen, J. (1994). Usability inspection methods. In *Conference Companion on Human Factors in Computing Systems* (pp. 413–414). New York, NY, USA: ACM. doi:10.1145/259963.260531

Nielsen, J. (1995). *Advances in Human-computer Interaction*. Intellect Books.

Nikora, L. W., Karapu, R., Hickey, H., & Te Awekotuku, N. (2004). Disabled Māori and disability support options. *University of Waikato Research Commons*. Retrieved from http://researchcommons.waikato.ac.nz/handle/10289/460

Norman, D. (1988). *The Psychology Of Everyday Things*. Basic Books. Retrieved from http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0465067093

Norman, D. A. (1983a). Design principles for human-computer interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1–10). New York, NY, USA: ACM. doi:10.1145/800045.801571

Norman, D. A. (1983b). Design rules based on analyses of human error. *Commun. ACM*, *26*(4), 254–258. doi:10.1145/2163.358092

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64. doi:10.1016/0010-0285(75)90004-3

Noyes, J. M., & Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics*, *50*(4), 514–519. doi:10.1080/00140130701235232

Nussbaum, M. (2007). Frontiers of Justice. Disability, Nationality, Species Membership. *Scandinavian Journal of Disability Research*, *9*(2), 133–136. doi:10.1080/15017410601003171

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, *43*(5), 971–990. doi:10.1016/j.paid.2007.03.017

Ojasalo, J., Seppälä, H., Suomalainen, N., & Moonen, R. (2010). Better technologies and services for smart homes of disabled people: Empirical findings from an explorative study among intellectually disabled. In *2010 2nd International Conference on Software Technology and Engineering (ICSTE)* (Vol. 1, pp. V1–251–V1–259). doi:10.1109/ICSTE.2010.5608845

Oliver, M., & Sapey, B. (2006). *Social Work with Disabled People, Third Edition*. PALGRAVE USA.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. doi:10.1037/h0043424

PAAS, F., RENKL, A., & SWELLER, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science*, *32*(1/2), 1–8.

Paas, F., Juhani E.Tabbers, HuibVan Gerven, Pascal W. M. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71.

Page, R. M. (1962). Man-Machine Coupling-2012 A.D. *Proceedings of the IRE*, *50*(5), 613–614. doi:10.1109/JRPROC.1962.288048

Paley, J. (2002). The Cartesian melodrama in nursing. *Nursing Philosophy*, *3*(3), 189–192. doi:10.1046/j.1466-769X.2002.00113.x

Palmeri, J. (2006). Disability Studies, Cultural Analysis, and the Critical Practice of Technical Communication Pedagogy. *Technical Communication Quarterly*, *15*(1), 49–65. doi:10.1207/s15427625tcq1501_5

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA, US: Academic Press.

Perenboom, R. J. M., & Chorus, A. M. J. (2003). Measuring participation according to the International Classification of Functioning, Disability and Health (ICF). *Disability & Rehabilitation*, *25*(11-12), 577–587. doi:10.1080/0963828031000137081

Petrie, H., Hamilton, F., King, N., & Pavan, P. (2006). Remote Usability Evaluations With Disabled People. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1133–1141). New York, NY, USA: ACM. doi:10.1145/1124772.1124942

Picard, R. W. (1995). *Affective Computing*. MIT Media Laboratory Perceptual Computing Section.

Picard, R. W. (2000). *Affective computing*. MIT Press.

Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, *36*(5), 741–773. doi:10.1016/0020-7373(92)90039-N

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. doi:10.1037/a0014996

Powers, W. T. (1973). *Behavior: The control of perception* (Vol. xi). Oxford, England: Aldine.

Pullin, G. (2009). *Design Meets Disability*. MIT Press.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*(2), 93–134. doi:10.1037/0033-295X.88.2.93

R A Rescorla, A. R. W. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory 2*.

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents*, *4*(2–4), 311–333. doi:10.1016/0376-6349(82)90041-4

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC-13*(3), 257 –266. doi:10.1109/TSMC.1983.6313160

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science*, *27*(2–3), 183–213. doi:10.1016/S0925-7535(97)00052-0

Reason, J. (1990a). *Human Error*. Cambridge University Press.

Reason, J. (1990b). The Contribution of Latent Human Failures to the Breakdown of Complex Systems. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), 475–484. doi:10.1098/rstb.1990.0090

Reason, J. (1995). Understanding adverse events: human factors. *Quality in Health Care*, *4*(2), 80–89. doi:10.1136/qshc.4.2.80

Reason, J. (2000). Human error: models and management. *BMJ*, *320*(7237), 768–770. doi:10.1136/bmj.320.7237.768

Reed, G. M., Lux, J. B., Bufka, L. F., Trask, C., Peterson, D. B., Stark, S., … Hawley, J. A. (2005). Operationalizing the International Classification of Functioning, Disability and Health in Clinical Settings. *Rehabilitation Psychology*, *50*(2), 122–131. doi:10.1037/0090-5550.50.2.122

Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. In Peter A. Hancock and Najmedin Meshkati (Ed.), *Advances in Psychology* (Vol. Volume 52, pp. 185–218). North-Holland. Retrieved from http://www.sciencedirect.com/science/article/pii/S0166411508623870

Rentsch, R., P, B., I, D. N., C, W., H, H., E, F., … I, B. (2003). The implementation of the “International Classification of Functioning, Disability and Health” (ICF) in daily practice of neurorehabilitation: an interdisciplinary project at the Kantonsspital of Lucerne, Switzerland. *Disability & Rehabilitation*, *25*(8), 411–421. doi:10.1080/0963828031000069717

Ritter, F. E., Baxter, G. D., & Churchill, E. F. (2014). *Foundations for Designing User-Centered Systems: What System Designers Need to Know about People*. Springer London.

Roberts, R. D., Beh, H. C., & Stankov, L. (1988). Hick’s law, competing-task performance, and intelligence. *Intelligence*, *12*(2), 111–130. doi:10.1016/0160-2896(88)90011-6

Rosander, P., Bäckström, M., & Stenberg, G. (2011). Personality traits and general intelligence as predictors of academic performance: A structural equation modelling approach. *Learning and Individual Differences*, *21*(5), 590–596. doi:10.1016/j.lindif.2011.04.004

Rose, F. D., Brooks, B. M., & Attree, E. A. (2002). An exploratory investigation into the usability and usefulness of training people with learning disabilities in a virtual environment. *Disability and Rehabilitation*, *24*(11-12), 627–633. doi:10.1080/09638280110111405

Rosenbaum, P., & Stewart, D. (2004). The world health organization international classification of functioning, disability, and health: a model to guide clinical thinking, practice and research in the field of cerebral palsy. *Seminars in Pediatric Neurology*, *11*(1), 5–10. doi:10.1016/j.spen.2004.01.002

Rothwell, J. C., Traub, M. M., Day, B. L., Obeso, J. A., Thomas, P. K., & Marsden, C. D. (1982). Manual Motor Performance in a Deafferented Man. *Brain*, *105*(3), 515–542. doi:10.1093/brain/105.3.515

Rowan, M., Gregor, P., Sloan, D., & Booth, P. (2000). Evaluating Web Resources for Disability Access. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies* (pp. 80–84). New York, NY, USA: ACM. doi:10.1145/354324.354346

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, *53*(1), 61–86. doi:10.1111/j.1464-0597.2004.00161

Rumelhart, D. E. (1977). *Introduction to human information processing*. Wiley.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507–514. doi:10.1016/j.neurobiolaging.2008.09.023

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive Functioning as a Potential Mediator of Age-Related Cognitive Decline in Normal Adults. *Journal of Experimental Psychology: General*, *132*(4), 566–594. doi:10.1037/0096-3445.132.4.566

Sauro, J. (2011). *A Practical Guide to the System Usability Scale: Background, Benchmarks and Best Practices*. CreateSpace.

Sauro, J., & Dumas, J. S. (2009). Comparison of Three One-question, Post-task Usability Questionnaires. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1599–1608). New York, NY, USA: ACM. doi:10.1145/1518701.1518946

Sauro, J., & Lewis, J. R. (2009). Correlations Among Prototypical Usability Metrics: Evidence for the Construct of Usability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1609–1618). New York, NY, USA: ACM. doi:10.1145/1518701.1518947

Schneider, D. W., & Anderson, J. R. (2011). A memory-based model of Hick's law. *Cognitive Psychology*, *62*(3), 193–222. doi:10.1016/j.cogpsych.2010.11.001

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. doi:10.1037/0033-295X.84.1.1

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599.

Scott Kelso, J. A. (1977). Planning and Efferent Components in the Coding of Movement. *Journal of Motor Behavior*, *9*(1), 33–47. doi:10.1080/00222895.1977.10735092

Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal*, *14*(2), 159–178. doi:10.1007/s11219-006-7600-8

Segregation, U. of the P. I. A. (1976). *The Union of the Physically Impaired Against Segregation and the Disability Alliance Discuss Fundamental Principles of Disability: Being a Summary of the Discussion Held on 22nd November, 1975 and Containing Commentaries from Each Organisation*. Union of the Physically Impaired Against Segregation.

Selwyn, N. (2009). The digital native – myth and reality. *Aslib Proceedings*, *61*(4), 364–379. doi:10.1108/00012530910973776

Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience*, *14*(5), 3208–3224.

SHAKESPEARE, T., & WATSON, N. (1997). Defending the Social Model. *Disability & Society*, *12*(2), 293–300. doi:10.1080/09687599727380

Shakespeare, T., & Watson, N. (2001). The social model of disability: An outdated ideology? *Research in Social Science and Disability*, *2*, 9–28. doi:10.1016/S1479-3547(01)80018-X

Shannon, C. E. (2001). A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, *5*(1), 3–55. doi:10.1145/584091.584093

Shannon, C. E., & Weaver, W. (1962). *The mathematical theory of communication. Univer*. of Illinois Press, Urbana.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. doi:10.1037/0033-295X.84.2.127

Shneiderman, B. (2000). Universal Usability. *Commun. ACM*, *43*(5), 84–91. doi:10.1145/332833.332843

Shorrock, S. T., & Kirwan, B. (2002). Development and application of a human error identification tool for air traffic control. *Applied Ergonomics*, *33*(4), 319–336. doi:10.1016/S0003-6870(02)00010-8

Shors, T. J., Anderson, M. L., Curlik II, D. M., & Nokia, M. S. (2012). Use it or lose it: How neurogenesis keeps the brain fit for learning. *Behavioural Brain Research*, *227*(2), 450–458. doi:10.1016/j.bbr.2011.04.023

Simeonsson, R. J., Lollar, D., Hollowell, J., & Adams, M. (2000). Revision of the International Classification of Impairments, Disabilities, and Handicaps: Developmental issues. *Journal of Clinical Epidemiology*, *53*(2), 113–124. doi:10.1016/S0895-4356(99)00133-X

Smith, M. A., & Shadmehr, R. (2005). Intact Ability to Learn Internal Models of Arm Dynamics in Huntington's Disease But Not Cerebellar Degeneration. *Journal of Neurophysiology*, *93*(5), 2809–2821. doi:10.1152/jn.00943.2004

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*(6), 482–489. doi:10.1037/h0055479

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics*, *37*(1), 55–79. doi:10.1016/j.apergo.2005.06.003

Stanton, N. A., & Baber, C. (2002). Error by design: methods for predicting device usability. *Design Studies*, *23*(4), 363–384. doi:10.1016/S0142-694X(01)00032-1

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science*, *47*(2), 227–237. doi:10.1016/j.ssci.2008.03.006

Stanton, P. N. A., Walker, D. G. H., & Salmon, A. P. P. M. (2013). *Human Factors Methods: A Practical Guide for Engineering and Design*. Ashgate Publishing, Ltd.

Stephen Hawking gets a tech upgrade from Intel, but keeps his original voice | News | Geek.com. (2014, February 12). Retrieved from http://www.geek.com/news/stephen-hawking-gets-a-tech-upgrade-from-intel-but-keeps-his-original-voice-1610726/

Stewart, M. G. (1992). Modelling human error rates for human reliability analysis of a structural design task. *Reliability Engineering & System Safety*, *36*(2), 171–180. doi:10.1016/0951-8320(92)90097-5

Sugeno, M., & Yasukawa, T. (1993, February). A fuzzy-logic-based approach to qualitative modeling. Retrieved August 6, 2013, from http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=390281

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. doi:10.1016/0364-0213(88)90023-7

Sweller, J., & Chandler, P. (1991). Evidence for Cognitive Load Theory. *Cognition and Instruction*, *8*(4), 351–362.

Sweller, J., van Merrienboer, J., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. doi:10.1023/A:1022193728205

Tapscott, D. (2008). *Grown Up Digital: How the Net Generation is Changing Your World*. McGraw Hill Professional.

Teasdale, N., Forget, R., Bard, C., Paillard, J., Fleury, M., & Lamarre, Y. (1993). The role of proprioceptive information for the production of isometric forces and for handwriting tasks. *Acta Psychologica*, *82*(1–3), 179–191. doi:10.1016/0001-6918(93)90011-F

Technology Vision 2014: Key IT Trends - Accenture. (2014). Retrieved August 24, 2014, from http://www.accenture.com/microsites/it-technology-trends-2014/Pages/tech-vision-report.aspx

Tedesco, D., & Tullis, T. (2006). A comparison of methods for eliciting post-task subjective ratings in usability testing. *Usability Professionals Association (UPA)*,*2006*, 1-9.

Terzi, L. (2004). The Social Model of Disability: A Philosophical Critique. *Journal of Applied Philosophy*, *21*(2), 141–157. doi:10.1111/j.0264-3758.2004.00269.x

Theofanos, M. F., & Redish, J. (Ginny). (2003). Bridging the Gap: Between Accessibility and Usability. *Interactions*, *10*(6), 36–51. doi:10.1145/947226.947227

ThinkPad X Series. (2014, November 30). In *Wikipedia, the free encyclopedia*. Retrieved from http://en.wikipedia.org/w/index.php?title=ThinkPad_X_Series&oldid=629492965

Thoroughman, K. A., & Shadmehr, R. (1999). Electromyographic Correlates of Learning an Internal Model of Reaching Movements. *The Journal of Neuroscience*, *19*(19), 8573–8588.

TSANG, P. S., & VELAZQUEZ, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, *39*(3), 358–381. doi:10.1080/00140139608964470

Tseng, Y., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of Neurophysiology*, *98*(1), 54–62. doi:10.1152/jn.00266.2007

Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick's Law in a Stochastic Race Model with Speed–Accuracy Tradeoff. *Journal of Mathematical Psychology*, *46*(6), 704–715. doi:10.1006/jmps.2002.1420

üStüN, T. B., Chatterji, S., Bickenbach, J., Kostanjsek, N., & Schneider, M. (2003). The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health. *Disability & Rehabilitation*, *25*(11-12), 565–571. doi:10.1080/0963828031000137063

Van Der Molen, H. T., Schmidt, H. G., & Kruisman, G. (2007). Personality characteristics of engineers. *European Journal of Engineering Education*, *32*(5), 495–501. doi:10.1080/03043790701433111

Vella, F., Vigouroux, N., & Gorce, P. (2009). Fitts's law interpretation as ergonomic metrics for motor impaired users. *Computer Methods in Biomechanics and Biomedical Engineering*, *12*(sup1), 249–250. doi:10.1080/10255840903097871

Vicente, K. J., & Rasmussen, J. (1990). The Ecology of Human-Machine Systems II: Mediating "Direct Perception" in Complex Work Domains. *Ecological Psychology*, *2*(3), 207–249. doi:10.1207/s15326969eco0203_2

Vigo, M., & Harper, S. (2013). Coping tactics employed by visually disabled users on the web. *International Journal of Human-Computer Studies*, *71*(11), 1013–1025. doi:10.1016/j.ijhcs.2013.08.002

Virzi, R. A. (1992). Refining the Test Phase of Usability Evaluation: How Many Subjects is Enough? *Hum. Factors*, *34*(4), 457–468.

Volkow, N. D., Gur, R. C., Gene-Jack, W., Fowler, J. S., & al, et. (1998). Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. *The American Journal of Psychiatry*, *155*(3), 344–9.

Waller, S. D., Langdon, P. M., & Clarkson, P. J. (2010). Using disability data to estimate design exclusion. *Universal Access in the Information Society*, *9*(3), 195–207. doi:10.1007/s10209-009-0168-x

Ward, A., & Mann, T. (2000). Don't mind if I do: Disinhibited eating under cognitive load. *Journal of Personality and Social Psychology*, *78*(4), 753–763. doi:10.1037/0022-3514.78.4.753

Warren, R., & Wertheim, A. H. (1990). *Perception & Control of Self-motion*. Psychology Press.

Weber, S. J., & Cook, T. D. (1972). Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychological Bulletin*, *77*(4), 273–295. doi:10.1037/h0032351

WHO | International Classification of Functioning, Disability and Health (ICF). (2013). Retrieved August 29, 2013, from http://www.who.int/classifications/icf/en/

WHO | International Classification of Functioning, Disability and Health (ICF). (2014). Retrieved September 23, 2013, from http://www.who.int/classifications/icf/en/

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Prentice Hall.

Windows 7 editions. (2014, December 3). In *Wikipedia, the free encyclopedia*. Retrieved from http://en.wikipedia.org/w/index.php?title=Windows_7_editions&oldid=631773129

Winkler, I., & Czigler, I. (1998). Mismatch negativity: deviance detection or the maintenance of the "standard." *NeuroReport*, *9*(17), 3809–3813. doi:10.1097/00001756-199812010-00008

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, *742*(1–2), 239–252. doi:10.1016/S0006-8993(96)01008-6

Wobbrock, J. O., Cutrell, E., Harada, S., & MacKenzie, I. S. (2008). An error model for pointing based on Fitts' law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1613–1622). New York, NY, USA: ACM. doi:10.1145/1357054.1357306

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 593–602. doi:10.1098/rstb.2002.1238

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An Internal Model for Sensorimotor Integration. *Science*, *269*(5232), 1880–1882. doi:10.2307/2889276